

Automatic sleep stages classification based on iterative filtering of electroencephalogram signals

Rajeev Sharma¹ · Ram Bilas Pachori¹ · Abhay Upadhyay¹

Received: 4 October 2016 / Accepted: 15 February 2017 / Published online: 3 March 2017
© The Natural Computing Applications Forum 2017

Abstract Computer-aided sleep monitoring system can effectively reduce the burden of experts in analyzing the large volume of electroencephalogram (EEG) recordings corresponding to sleep stages. In this paper, a new technique for automated classification of sleep stages based on iterative filtering of EEG signals is presented. In order to perform sleep stages classification, the EEG signals are decomposed using iterative filtering method. The modes obtained from iterative filtering of EEG signal can be considered as amplitude-modulated and frequency-modulated (AM-FM) components. The discrete energy separation algorithm (DESA) is applied to the modes to determine amplitude envelope and instantaneous frequency functions. The extracted amplitude envelope and instantaneous frequency functions have been used to compute Poincaré plot descriptors and statistical measures. The Poincaré plot descriptors and statistical measures are applied as input features for different classifiers in order to classify sleep stages. The classifiers namely, naïve Bayes, k -nearest neighbor, multilayer perceptron, C4.5 decision tree, and random forest are applied in order to classify the EEG epochs corresponding to various sleep stages. The experimental study has been performed on online available Sleep-EDF database for two-class to six-class classification of sleep stages based on EEG signals. The two-class to six-

class classification problems are formulated by taking different combinations of EEG signals corresponding to various sleep stages. The comparison of the results is presented for different multi-class classification problems with the other recently proposed methods. The results show that the proposed method has provided better tenfold cross-validation classification accuracy than other existing methods.

Keywords Sleep stages · Electroencephalogram signal · Iterative filtering · Teager energy operator · Classifiers

1 Introduction

Sleep analysis is an important area of research due to its importance for the study of disorders related to sleep [25]. In a healthy human brain, several psychophysiological states occur during sleep, referred as sleep stages [66]. Sleep-related disorders like sleep apnea, insomnia, and narcolepsy affect the normal life of the patients; therefore, effective diagnosis is very useful for such patients. Rechtschaffen and Kales (R&K) developed a standard criterion to manually score the sleep stages using the electroencephalogram (EEG) signals [58]. Based on the R&K standard, there are six sleep stages categorized in four non-rapid eye movement (NREM) sleep stage 1 (S1), NREM sleep stage 2 (S2), NREM sleep stage 3 (S3), NREM sleep stage 4 (S4), followed by rapid eye movement (REM) sleep stage and remaining stage as wakefulness or awake (W) stage. Recently, American Academy of Sleep Medicine (AASM) proposed a new criterion for scoring the sleep stages. The AASM standard suggests merging the S3 and S4 stages into a single sleep stage. According to this standard, the NREM stage comprises of only three sleep

✉ Rajeev Sharma
phd1301102007@iiti.ac.in
Ram Bilas Pachori
pachori@iiti.ac.in
Abhay Upadhyay
phd1301102002@iiti.ac.in

¹ Discipline of Electrical Engineering, Indian Institute of Technology Indore, Indore 453552, India

stages [27]. Manual sleep EEG scoring takes cumbersome efforts to accomplish sleep EEG scoring due to the long length of recordings acquired during night. Therefore, computer-assisted automatic sleep scoring system can be more efficient for diagnosis of sleep-related disorders using EEG signals. Such type of system can be helpful for clinicians to speed up the process of diagnosis for the various sleep-related abnormalities.

Recently, several techniques are proposed in the literature to accomplish the task of designing automatic sleep scoring system based on polysomnographic (PSG) recordings [48]. The PSG recording is a collection of simultaneously recorded different physiological signals which include EEG, electromyogram (EMG), electrooculogram (EOG), electrocardiogram (ECG) etc. In Agarwal and Gotman [5], an automatic sleep stages classification method based on the segmentation and self-organization (clustering) techniques is presented. The features extracted from segmented PSG recordings are amplitude, dominant rhythms, frequency-weighted energy, alpha-slow-wave index etc., and the sleep stages classification is performed using *k*-mean clustering algorithm. In another study, a methodology based on adaptive neuro-fuzzy classifier is presented for sleep stages classification from PSG recorded from infants [24]. Krakovská and Mezeioá [39] analyzed the PSG data recorded from 20 healthy subjects and total of 74 measures have been extracted for sleep scoring. The sleep stages are classified based on the quadratic discriminant analysis.

The above-mentioned methods for automated classification of sleep stages are based on PSG which requires recording of a number of physiological signals and suffers from their own limitations. Several methodologies have been developed which use only EEG signals for detection of sleep stages. In Acharya et al. [3], for the sleep stages detection, 29 nonlinear dynamic measures are computed from EEG signals. The nonlinear dynamic features include the higher order spectra (HOS) and the recurrence quantification analysis (RQA)-based features. The HOS-based features address the nonlinearity and deviation from Gaussian nature present in the EEG signals [3]. Similarly, RQA-based features are used to quantify the irregularity present in the EEG signals. However, only significance of the features is evaluated using analysis of variance (ANOVA) test for the discrimination of the various sleep stages without applying any classification technique. Finally, the results are presented in terms of *F* value and *p* value obtained from the ANOVA test. Liang et al. [43] proposed a method which is based on multiscale entropy and autoregressive model for automatic sleep scoring using EEG signals. In this study, the authors only presented results for the six-class classification. They obtained the overall sensitivity (Sen) which is 76.9% for classification

of six classes (W, S1, S2, S3, S4, and REM). The energy features are proposed for classification of five sleep stages using single channel EEG signals [25]. In their study, in order to perform classification, the recurrent neural network classifier has been applied. They have only presented results for the five-class (W, S1, S2, S3–S4, REM) classification problem. The average classification accuracy obtained for five-class classification problem is 87.2%. Imtiaz and Rodriguez-Villegas [28] proposed a feature based on the spectral edge frequency (SEF) related to the frequency band of 8–16 Hz and combined it with absolute power and relative power features extracted from EEG signals for REM stage detection. In this work, the authors only performed the REM stage detection and the obtained Sen for the REM sleep detection which is 83%. In Tsinalis et al. [64], for the classification of the sleep stages, a stacked sparse autoencoder-based technique is presented. They obtained the mean classification accuracy for individual sleep stages as 84% with a range of 82–86%. The authors performed the classification using cross-validation with 20-fold for computing the performance of the classifier. They also addressed the problem of class imbalance using the random sampling technique. However, they only considered the five-class classification problem in their study. Time–frequency image of EEG signals obtained using smoothed pseudo Wigner–Ville distribution is suggested by Bajaj and Pachori [9] for sleep stages classification. The obtained time–frequency images of EEG signals are segmented according to the different EEG rhythms and features are computed from the histogram of these segmented images. Finally, multi-class support vector machine (SVM) is used for sleep stages classification using EEG signals. The method presented in Bajaj and Pachori [9] has been studied only for six-class and five-class classification problems.

An approach based on the horizontal visibility graph (HVG) algorithm is presented by Zhu et al. [73] for sleep stages analysis based on EEG signals. They used mean degree and degree distribution features extracted from single channel EEG signals with SVM as a classifier for classification of sleep stages. The authors performed the classification for six-class, five-class, four-class (W, S1–S2, S3–S4, and REM), three-class (W, NREM, and REM), and two-class (W, NREM–REM) classification problems. The obtained classification accuracies for these classification problems are 97.90, 92.60, 89.30, 88.90 and 87.50 %, respectively. They divided the half of the epochs for training and remaining half for the testing purpose. Recently, Hassan and Bhuiyan [23] used complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and bootstrap aggregating techniques for automatic sleep staging based on EEG signals. They computed the higher order statistical moments from

intrinsic mode functions (IMFs) extracted by ensemble empirical mode decomposition (EEMD) process and used as features with the bootstrap aggregating algorithm for classification of various sleep stages using EEG signals. In this work, classification is performed for six-class to two-class classification problems. The reported classification accuracies for these classification problems are 99.48, 94.10, 92.14, 90.69 and 86.89%, respectively. A new method [22] employed recently proposed tunable Q wavelet transform (TQWT) for decomposing the sleep EEG epochs. The spectral features computed from the TQWT subbands have been used as an input to the random forest classifier. They obtained the accuracies for six-class to two-class classification problems as 97.50, 94.80, 92.11, 91.50 and 90.38%, respectively. Ronzhina et al. [59] classified sleep EEG epochs corresponding to Pz–Oz channel and presented results for various multi-class sleep stages classification with different architectures of artificial neural network (ANN) classifier. The obtained classification accuracy values for six-class, four-class, three-class, and two-class classification problems are 76.70, 81.55, 90.31 and 98.62%, respectively. In this study, tenfold cross-validation is applied to compute the classification accuracy for the stated classification problems. In another method, the different sleep stages are classified using autoregressive coefficients computed from the epochs of the EEG signals [34]. They proposed a fast classification method based on the partial least squares (PLS) regression classifier. They analyzed the different length of epochs using their methodology for classification of awake and sleep stages. They performed tenfold cross-validation for classification and presented patient-specific results.

The literature review presented in previous paragraphs shows that most of the studies targeted their methodology only for six-class and/or five-class classification problems. The other classification problems corresponding to four-class, three-class, and two-class have not been addressed. In several methods [22, 23], all the classification problems are considered; however, they used the pre-partitioned training and testing data. The classification performance is not evaluated for any type of cross-validation. The single split of training and testing of data may be biased for the data selected for training. On the other

hand, the cross-validation procedure is less sensitive for the partitioning of the data. In addition, it can also be observed from the literature review that there is still scope for improving the classification performance for the various multi-class classification problems with cross-validation approach.

In this article, we present a method for automatic classification of sleep stages from EEG signals. Our objective is to study the features extracted from the amplitude envelope and instantaneous frequency functions of the modes obtained using iterative filtering method for classification of sleep stages. The block diagram which outlines the main steps of the proposed method is presented in Fig. 1. Firstly, we applied the iterative filtering method presented by Cicone et al. [15] which uses the smooth filters with compact support for extraction of different modes. Iterative filtering [15, 44] is an alternative algorithm for the empirical mode decomposition (EMD) method [26], and it can better handle the issues related to the sifting process and cubic spline algorithm. In addition, as compared to EMD method, iterative filtering method has better stability [67]. Moreover, the iterative filtering method is more suitable for horizontal comparison which means the modes of the same index belonging to different signals contain comparable information [67]. Furthermore, the performance of the iterative filtering is yet to be explored for sleep stages classification from EEG signals. During the span of different sleep stages, the amplitude and frequency contents of EEG signals change [20]; therefore, obtaining the amplitude envelope and instantaneous frequency functions from the modes obtained using iterative filtering can lead to effective feature computation. The discrete energy separation algorithm (DESA) [47] is applied to extract amplitude envelope and instantaneous frequency functions from the modes. For the purpose of sleep stages classification, the features have been extracted from the amplitude envelope and instantaneous frequency functions computed from the different modes. The studied features are Poincaré plot descriptors and higher order statistical moments. Finally, naïve Bayes, k -nearest neighbor, C4.5 decision tree classifier, multilayer perceptron, and random forest classifiers are studied for classification of different sleep stages. The proposed methodology

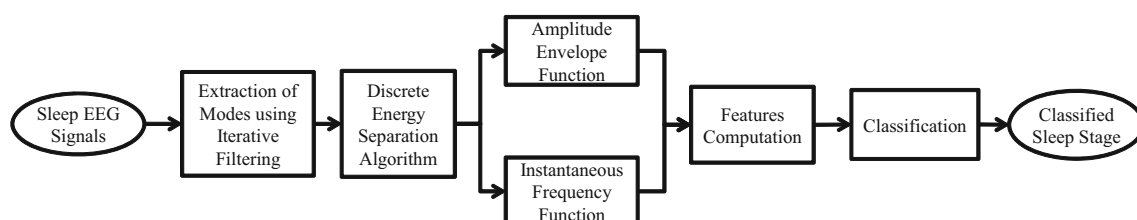


Fig. 1 Block diagram of proposed methodology

is evaluated for all the studied different multi-class classification problems. The tenfold cross-validation procedure is employed to evaluate the classifiers' performance for these multi-class classification problems.

The organization of remaining paper is as follows: Sect. 2 gives the brief description of the studied sleep EEG dataset, iterative filtering method, DESA method, and different classifiers applied for classification purpose. In Sect. 3, results obtained using proposed methodology are presented and discussed in Sect. 4. Finally, the conclusion is presented in Sect. 5.

2 Dataset and methods

2.1 Dataset

The sleep EEG recordings which are used to evaluate the proposed methodology have been taken from Sleep-EDF database available online at PhysioNet website [19, 36]. The recordings obtained from eight subjects have been used for performing the experimental study. The EEG recordings used in this paper are denoted as SC4002E0, SC4012E0, SC4102E0, SC4112E0, ST7022J0, ST7052J0, ST7121J0, and ST7132J0. The first four recordings which are indicated as SC were obtained from the healthy volunteers. On the other hand, last four recordings which are indicated as ST were obtained from the subjects having mild difficulty in falling asleep [36, 73]. In this database, each PSG recording has two EEG signals recorded from Fpz–Cz and Pz–Oz channels, one EOG signal, one EMG signal, and one oro-nasal respiration signal. The EEG signals are sampled at the sampling rate of 100 Hz. In this study, the signals recorded from Pz–Oz channel are selected for evaluation of proposed methodology as it has been used for evaluation of the previously developed methodologies for detection of sleep stages [23, 43, 73].

The EEG sleep recordings were annotated by experts. These annotations have been used in this work as a reference in order to verify obtained results from the proposed methodology. The annotations have been given separately in the hypnogram files which are available in the database. The sleep annotations are provided for every 30 s of epoch for each EEG signal. According to R&K standard, each epoch is annotated as one out of following eight classes: Awake (W), S1, S2, S3, S4, REM, movement time (M) and unknown state (U). The tenfold cross-validation [38] is used for assessment of the classification performance of the studied classifiers. The number of epochs for each analyzed class is listed in Table 1. In Fig. 2, EEG epochs corresponding to different sleep stages are depicted. It should be noted that the epochs related to the M and U states are not included in the experimental study. According to R&K

Table 1 Number of epochs considered for each sleep condition from Sleep-EDF database

Sleep condition	Total number of selected epochs
Awake condition	8003
S1 sleep stage	604
S2 sleep stage	3621
S3 sleep stage	672
S4 sleep stage	627
REM sleep stage	1609
All conditions	15,136

standard, the sleep cycles are distributed across six sleep stages. The six-class classification case is clinically relevant and frequently studied in various methods proposed in the literature [23, 43, 73]. According to AASM standard, it has been proposed to combine the S3 and S4 sleep stages together to form a single class. Therefore, this newly proposed standard suggests the importance of the five-class classification problem. Hence, the five-class classification is considered for sleep stages classification. The S1 and S2 stages can be combined into a single class as both are called shallow sleep [23, 70], which leads to four-class classification case. The awake, REM, and non-REM (S1, S2, S3, and S4) sleep stages detection can be useful for some clinical diagnosis of the sleep disorders like REM sleep behavior disorder [29, 30]. Therefore, three-class classification problem is also studied in this work. It can be useful to differentiate the sleep recordings from those in awake conditions; hence, the two-class classification is also employed in the experiments.

2.2 Iterative filtering-based decomposition

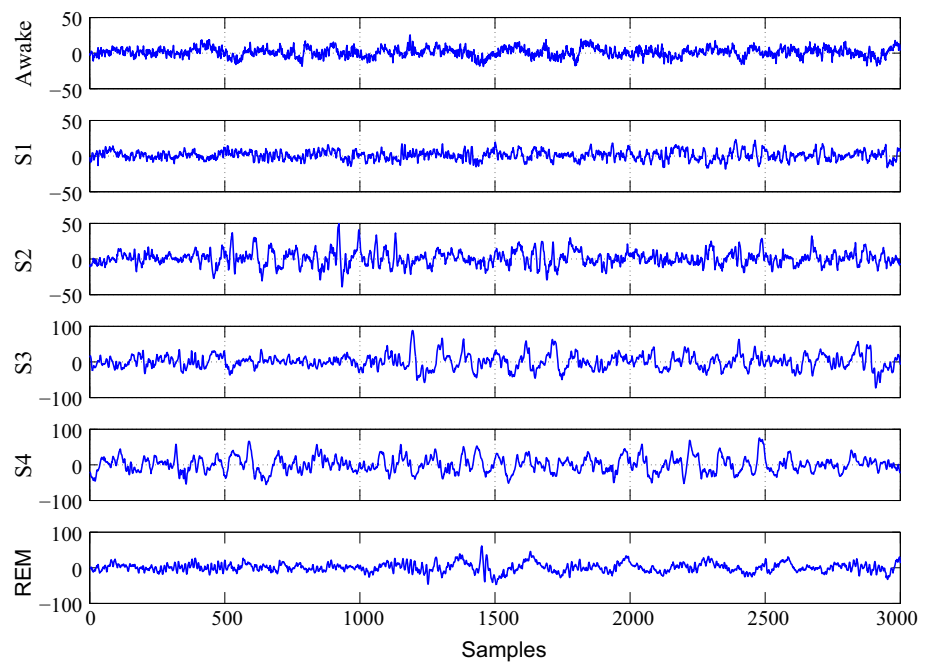
Iterative filtering is an iterative approach for the decomposition of the nonlinear and nonstationary signals [44, 68]. The obtained modes from the iterative filtering satisfy the conditions of intrinsic mode functions (IMFs) as mentioned in Huang et al. [26] and Cicone et al. [15]. The iterative filtering uses low pass filters with modified sifting process in order to obtain a mode [67]. The method of extracting modes using iterative filtering of a signal can be explained briefly as follows [15]:

For a given signal $g(t)$, an operator L which determines moving average of $g(t)$ can be defined as:

$$L[g(t)] = \int_{-l}^{+l} g(t + \tau)h(\tau)dt. \quad (1)$$

The function $h(t)$ is a filter function, and l denotes the mask length [15]. Another operator F , with initialization $g_1 = g$, can be defined as: $F_{1,n}(g_n) = g_n - L_n^{(1)}(g_n) = g_{n+1}$. The operator $F_{1,n}$ captures the fluctuation part from $g_n(t)$.

Fig. 2 Plots of EEG signals corresponding to different EEG sleep stages (the unit of amplitude is in μV)



In order to extract the first mode M_1 , the $F_{1,n}$ is applied on the signal as: $M_1 = \lim_{n \rightarrow \infty} F_{1,n}(g_n)$. Here, filtering operator $L_n^{(1)}$ has mask length l_n for every iteration number n and superscript denotes the first mode.

Further, to extract next mode, operator F can be applied to remainder signal $g - M_1$. This process of extracting modes can be repeated until remainder has at most one local extremum. For more detail information about the iterative filtering and its algorithm, the paper by Cicone et al. [15] can be referred.

In order to perform iterative filtering, we have used the implementation proposed by Cicone et al. [15]. In this work, the authors used Fokker–Planck equation to design low pass filter with compact support. This filter has an advantage of being smooth enough that it vanishes to zero at both ends. Consequently, these filters do not produce artificial oscillations as in the case of double average filters [15]. The MATLAB codes for implementation of the iterative filtering using Fokker–Planck filters are available at <http://www.mathworks.com/matlabcentral/fileexchange/53405-iterative-filters>. The iterative filtering process consists of two loops, one as an inner loop and other as an outer loop. The mask length can be updated at each step of the inner loop. In the implemented algorithm, the mask length l_n is computed for the first step and the same value is used for remaining steps. The length of the mask can be computed as [15]: $l_n = 2 \lfloor \lambda \frac{N}{\kappa} \rfloor$, where N denotes the length of the signal, κ is the number of extreme points, λ is a constant with value 1.6, and $\lfloor \cdot \rfloor$ indicates greatest integer operator. An EEG signal corresponding to awake stage

and its first seven modes ($M_1 - M_7$) obtained using iterative filtering are shown in Fig. 3.

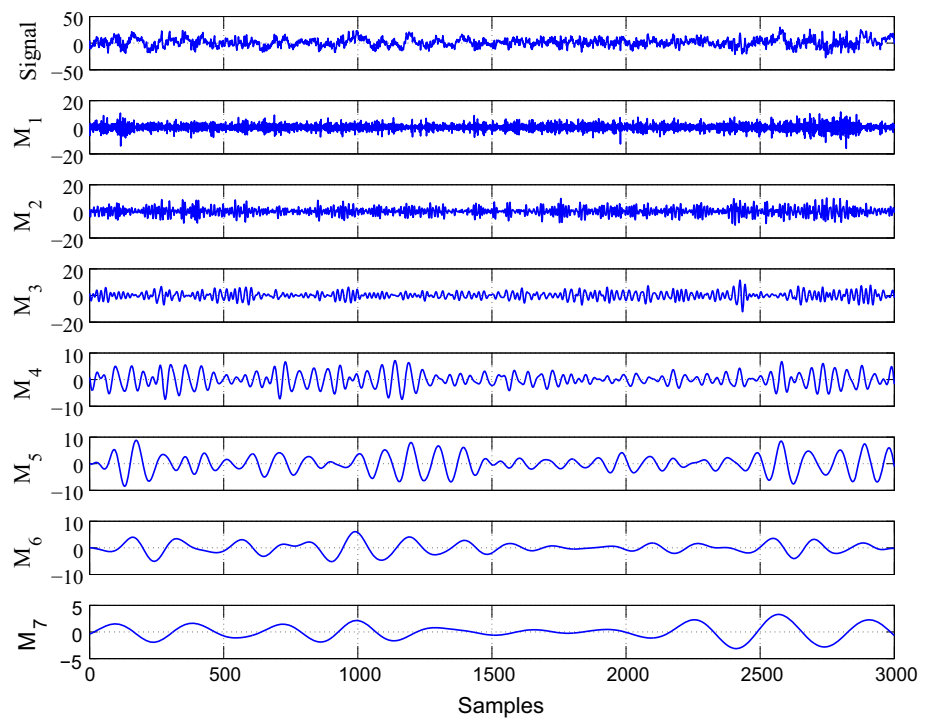
2.3 Discrete energy separation algorithm (DESA)

The variation in the amplitude and frequency parameters can be useful for effective identification of different sleep stages using EEG signals as amplitude and frequency characteristics change across various sleep epochs [20]. These variations in amplitude and frequency can reflect in various modes of EEG signals and can be effectively captured by computing amplitude envelope and instantaneous frequency functions. The amplitude envelope and instantaneous frequency functions of the modes can be separated using the Hilbert transform (HT) and Teager energy operator (TEO) [56]. The HT is applied globally on the whole signal. On the other hand, TEO [32] is more suitable for local estimation of the amplitude envelope and instantaneous frequency functions [11, 53]. In addition, TEO is found useful for many signal processing applications [33, 50, 51, 61, 72]. It is a nonlinear operator, which can be computed for discrete-time signal $I(n)$ based on three samples as [32, 47, 53, 72]:

$$\Psi[I(n)] = I^2(n) - I(n+1)I(n-1). \tag{2}$$

In our case, $I(n)$ is a mode obtained from iterative filtering method. The obtained mode can be considered as an amplitude-modulated and frequency-modulated (AM-FM) signal. The TEO can be used to estimate the amplitude envelope and instantaneous frequency functions of an AM-FM signal [47, 53, 72]. The algorithm developed to separate the amplitude envelope and instantaneous frequency

Fig. 3 Awake stage EEG signal and its first seven obtained modes using iterative filtering (the unit of amplitude is in μV)



functions based on TEO is called discrete energy separation algorithm (DESA) [47]. The computational complexity of DESA is less than Hilbert transform separation algorithm [56]. Using DESA, following expressions can be used to compute amplitude envelope and instantaneous frequency functions [47]:

$$I_{\text{FM}}(n) \approx \arccos\left(1 - \frac{\Psi[H(n)] + \Psi[H(n+1)]}{4\Psi[I(n)]}\right) \quad (3)$$

$$I_{\text{AM}}(n) \approx \sqrt{\frac{\Psi[I(n)]}{\left[1 - \left(1 - \frac{\Psi[H(n)] + \Psi[H(n+1)]}{4\Psi[I(n)]}\right)^2\right]}} \quad (4)$$

where, $H(n)$ is a time-domain difference signal of $I(n)$ given by $H(n) = I(n) - I(n-1)$; $\Psi[\cdot]$ denotes the TEO; $I_{\text{FM}}(n)$ denotes instantaneous frequency function of $I(n)$; and $I_{\text{AM}}(n)$ indicates the amplitude envelope of $I(n)$ at sample n . The amplitude envelope $I_{\text{AM}}(n)$ and instantaneous frequency function $I_{\text{FM}}(n)$ computed using DESA technique for modes shown in Fig. 3 are shown in Fig. 4.

2.4 Feature extraction

Feature extraction is an important step for automatic classification of different sleep stages using EEG signals. This step extracts the characteristic patterns of EEG signals corresponding to different sleep stages. In this work, we have explored the features obtained from amplitude

envelope (I_{AM}) and instantaneous frequency (I_{FM}) functions of the modes from iterative filtering method. Ten different features from the I_{AM} and I_{FM} functions are computed for each mode of the EEG epoch. The Poincaré plot of the I_{AM} and I_{FM} functions is used for extracting geometric descriptors [13, 55]. These descriptors are the width of Poincaré plot, the length of Poincaré plot, the area of Poincaré plot, acceleration, and deceleration. From a given signal $s(n)$, two vectors U and V are formed which represent the points (u_i, v_i) on Poincaré plot with $i = 1, 2, \dots, N$. The vector V is one sample delayed version of U . Then following parameters can be defined as [13, 55]:

$$D_i^1 = \frac{|u_i - u_c - v_i + v_c|}{\sqrt{2}}$$

and

$$D_i^2 = \frac{|u_i - u_c + v_i - v_c|}{\sqrt{2}}.$$

The point (u_c, v_c) represents the centroid of the distribution of N points on Poincaré plot.

The measures of asymmetry of Poincaré plot about the line of identity are defined for RR interval signals in [55]. We also used these measures of asymmetry of Poincaré plot as the features for sleep stages classification. In order to compute these features, the following parameters are used [55]:

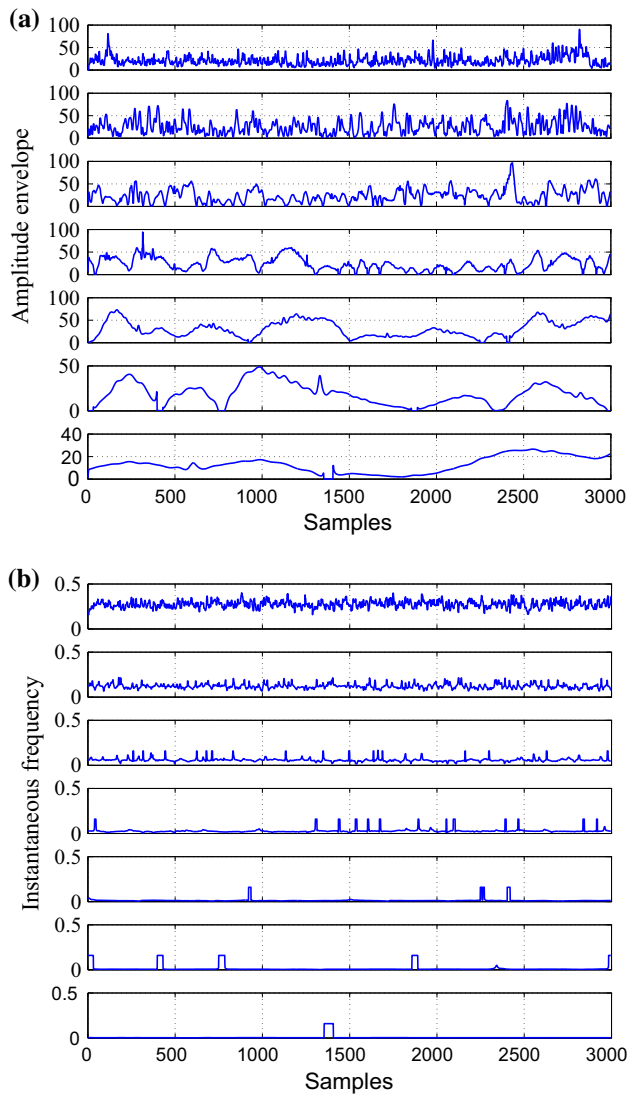


Fig. 4 Results of DESA applied on the modes obtained from iterative filtering of awake EEG signal: **a** amplitude envelope functions of different modes (amplitude is in μV) and **b** instantaneous frequency functions of different modes. Panels from *top* to *bottom* show plots corresponding to modes from M_1 to M_7

$$SD_{\text{up}} = \frac{1}{N} \sum_{i=1}^{n_U} (d_U^i)^2 \tag{5}$$

$$SD_{\text{down}} = \frac{1}{N} \sum_{i=1}^{n_D} (d_D^i)^2 \tag{6}$$

and

$$SD_I = SD_{\text{up}} + SD_{\text{down}}. \tag{7}$$

The variables n_U and n_D represent the number of points above and below the line of identity, respectively. The parameters d_U^i are the distance of i th point above the line of identity. Similarly, the distance of i th point below the line of identity is denoted by d_D^i . Apart from five features

Table 2 Features extracted from amplitude envelope of a mode obtained from iterative filtering method

Notation	Feature
F1	Width of Poincare plot [13, 55], $SD_1 = \frac{1}{N} \sum_{i=1}^N (D_i^1)^2$
F2	Length of Poincare plot [13, 55], $SD_2 = \frac{1}{N} \sum_{i=1}^N (D_i^2)^2$
F3	Area of Poincare plot [13, 55], $\text{Area}_p = \pi \times SD_1 \times SD_2$
F4	Acceleration [55], $C_{\text{up}} = \frac{SD_{\text{up}}^2}{SD_I^2}$
F5	Deceleration [55], $C_{\text{down}} = \frac{SD_{\text{down}}^2}{SD_I^2}$
F6	Mean, $\mu = \frac{1}{M} \sum_{j=1}^M w_j$
F7	Variance, $v^2 = \frac{1}{M} \sum_{j=1}^M (w_j - \mu)^2$
F8	Median
F9	Kurtosis = $\frac{1}{M} \sum_{j=1}^M \left(\frac{w_j - \mu}{v}\right)^4$
F10	Skewness = $\frac{1}{M} \sum_{j=1}^M \left(\frac{w_j - \mu}{v}\right)^3$

extracted from the Poincaré plot of I_{AM} , we also computed five statistical features as mean, variance, median, kurtosis, and skewness [37] from the modes. For a signal of M samples $I_{AM} = \{w_1, w_2, w_3, \dots, w_M\}$ all these extracted features computed for I_{AM} are listed and defined in Table 2. The features extracted from I_{AM} are indexed as F1–F10, respectively, and similarly, all ten features are extracted from I_{FM} and indexed by F11–F20, in exact same order.

The variation in the extracted features from the iterative filtering and EMD method is different. Table 3 reports the range values in terms of confidence intervals [37] of the mean of features extracted from iterative filtering and EMD method. In Table 3, the 99% confidence interval [4] of the features extracted from different features are presented. It should be noted that in order to obtain the confidence interval, the feature values less than 99 percentile of the feature are considered. The values of a feature corresponding to each class are fitted to the normal distribution. The obtained normal distribution is used to compute the confidence interval of the feature corresponding to each class.

The suitability of the features computed from iterative filtering method and EMD is examined using the statistical test. The Kruskal–Wallis statistical test [40] is a nonparametric test for multigroup data and a distribution-free alternative to the one-way ANOVA test [63]. It results in p value which measures the significance of Chi-squared statistics [63]. In Villanueva et al. [65], the p value resulted using Kruskal–Wallis statistical test is used for examining the significance of the features for multi-class comparison of mass spectrometry (MS)-based serum peptide profiling data. The data were obtained from three types of cancer patients and control group. Similarly, in the presented

Table 3 Confidence intervals [mean: (confidence interval = 99%)] and *p* values obtained using Kruskal–Wallis statistical test for various features extracted from first mode and IMF obtained from iterative filtering and EMD method, respectively, for various sleep stages

Feature	Method	Awake	S1	S2	S3	S4	REM	<i>p</i> value
F1	Iterative filtering	5.4825; (5.3832, 5.5818)	1.9744; (1.8259, 2.1230)	1.6369; (1.6082, 1.6657)	1.8716; (1.8102, 1.9330)	2.0135; (1.9654, 2.0616)	1.4536; (1.4105, 1.4966)	0
	EMD	5.4963; (5.3993, 5.5933)	1.9612; (1.8115, 2.1110)	1.6374; (1.6086, 1.6663)	1.7189; (1.6519, 1.7858)	1.6693; (1.6206, 1.7179)	1.4699; (1.4293, 1.5105)	0
F2	Iterative filtering	35.1979; (34.5378, 35.8580)	13.9896; (12.6690, 15.3102)	14.1137; (13.7715, 14.4559)	16.6559; (15.9512, 17.3605)	18.2466; (17.7094, 18.7837)	9.3079; (8.8570, 9.7589)	0
	EMD	39.0837; (38.3983, 39.7691)	18.1701; (16.8019, 19.5382)	23.3326; (22.8811, 23.7840)	23.5561; (22.5471, 24.5652)	22.9823; (22.2742, 23.6905)	15.3356; (14.7030, 15.9682)	0
F3	Iterative filtering	844.4969; (816.3248, 872.6690)	139.6875; (107.3409, 172.0341)	88.0605; (81.6815, 94.4395)	109.3077; (100.3771, 118.2383)	120.5899; (113.8334, 127.3464)	56.3277; (48.7233, 63.9321)	0
	EMD	910.5157; (881.2224, 939.8089)	161.1201; (128.9646, 193.2756)	138.3896; (131.6048, 145.1743)	144.4409; (132.9852, 155.8966)	128.6063; (120.8383, 136.3744)	87.2812; (79.5709, 94.9915)	0
F4	Iterative filtering	0.49663; (0.4961, 0.4971)	0.49682; (0.4954, 0.4983)	0.49724; (0.4967, 0.4978)	0.49693; (0.4958, 0.4981)	0.497; (0.4958, 0.4982)	0.49782; (0.4971, 0.4986)	0.4481
	EMD	0.49537; (0.4947, 0.4961)	0.49397; (0.4918, 0.4962)	0.49548; (0.4946, 0.4963)	0.49493; (0.4926, 0.4972)	0.49609; (0.4939, 0.4983)	0.49603; (0.4948, 0.4973)	0.4499
F5	Iterative filtering	0.5016; (0.5011, 0.5021)	0.5024; (0.5010, 0.5038)	0.5025; (0.5020, 0.5030)	0.5030; (0.5019, 0.5042)	0.5028; (0.5016, 0.5041)	0.5016; (0.5008, 0.5024)	0.0506
	EMD	0.5028; (0.5021, 0.5034)	0.5033; (0.5011, 0.5055)	0.5036; (0.5028, 0.5044)	0.5032; (0.5012, 0.5052)	0.5020; (0.4995, 0.5045)	0.5023; (0.5011, 0.5035)	0.4662
F6	Iterative filtering	39.2272; (38.5584, 39.8960)	16.2172; (15.3433, 17.0910)	16.4408; (16.1600, 16.7217)	19.0066; (18.3553, 19.6580)	20.9609; (20.4817, 21.4402)	12.8848; (12.5849, 13.1848)	0
	EMD	40.9737; (40.2924, 41.6551)	16.6797; (15.7827, 17.5767)	17.0453; (16.7380, 17.3525)	16.8020; (16.0563, 17.5477)	16.0304; (15.5616, 16.4993)	13.3486; (13.0015, 13.6958)	0
F7	Iterative filtering	899.2338; (868.8507, 929.6169)	179.0988; (135.0337, 223.1640)	133.1033; (124.3219, 141.8847)	165.6924; (150.9468, 180.4380)	182.2008; (170.6117, 193.7899)	69.2240; (56.9129, 81.5351)	0
	EMD	1063.7112; (1029.6395, 1097.7829)	252.2914; (203.9867, 300.5961)	329.3742; (316.1660, 342.5824)	330.4586; (303.3857, 357.5314)	289.1975; (271.5164, 306.8786)	167.3165; (150.4825, 184.1505)	0
F8	Iterative filtering	33.1901; (32.6200, 33.7601)	13.5558; (12.9767, 14.1350)	13.7082; (13.5004, 13.9161)	15.8015; (15.2680, 16.3350)	17.4936; (17.0977, 17.8895)	11.4890; (11.2552, 11.7228)	0
	EMD	34.8853; (34.2921, 35.4785)	13.1212; (12.5204, 13.7220)	11.0618; (10.8778, 11.2457)	10.8202; (10.3504, 11.2900)	10.3359; (10.0617, 10.6102)	10.3954; (10.1767, 10.6141)	0
F9	Iterative filtering	8.4334; (8.2734, 8.5935)	7.6695; (7.0764, 8.2626)	7.0564; (6.9129, 7.1999)	7.1011; (6.7089, 7.4933)	6.5038; (6.2056, 6.8020)	6.5504; (6.2521, 6.8488)	0
	EMD	7.8929; (7.7389, 8.0469)	10.8258; (10.0653, 11.5864)	10.2570; (10.0207, 10.4932)	11.4344; (10.8050, 12.0638)	11.7675; (11.1158, 12.4191)	12.1577; (11.6853, 12.6301)	0

Table 3 continued

Feature	Method	Awake	S1	S2	S3	S4	REM	p value
F10	Iterative filtering	1.7048; (1.6838, 1.7258)	1.5622; (1.4812, 1.6432)	1.6449; (1.6228, 1.6671)	1.6645; (1.6121, 1.7170)	1.5926; (1.5426, 1.6426)	1.3459; (1.3039, 1.3879)	0
	EMD	1.5880; (1.5661, 1.6100)	2.1063; (2.0053, 2.2072)	2.3534; (2.3245, 2.3824)	2.4911; (2.4201, 2.5621)	2.5920; (2.5213, 2.6628)	2.3027; (2.2444, 2.3610)	0
F11	Iterative filtering	0.0801; (0.0799, 0.0803)	0.0694; (0.0689, 0.0700)	0.0684; (0.0681, 0.0686)	0.0691; (0.0684, 0.0697)	0.0686; (0.0681, 0.0692)	0.0700; (0.0697, 0.0703)	0
	EMD	0.1516; (0.1512, 0.1520)	0.1250; (0.1235, 0.1265)	0.1234; (0.1226, 0.1241)	0.1280; (0.1261, 0.1300)	0.1261; (0.1250, 0.1273)	0.1275; (0.1264, 0.1286)	0
F12	Iterative filtering	0.3186; (0.3180, 0.3191)	0.3397; (0.3372, 0.3421)	0.3739; (0.3729, 0.3748)	0.3817; (0.3798, 0.3836)	0.3843; (0.3827, 0.3860)	0.3458; (0.3447, 0.3469)	0
	EMD	0.4962; (0.4953, 0.4972)	0.5164; (0.5116, 0.5213)	0.6053; (0.6033, 0.6072)	0.6254; (0.6212, 0.6296)	0.6264; (0.6232, 0.6297)	0.5505; (0.5470, 0.5541)	0
F13	Iterative filtering	0.0802; (0.0799, 0.0804)	0.0742; (0.0733, 0.0752)	0.0805; (0.0801, 0.0808)	0.0831; (0.0822, 0.0841)	0.0832; (0.0824, 0.0840)	0.0762; (0.0758, 0.0766)	0
	EMD	0.2367; (0.2360, 0.2374)	0.2039; (0.2003, 0.2075)	0.2336; (0.2318, 0.2354)	0.2501; (0.2458, 0.2544)	0.2487; (0.2459, 0.2514)	0.2225; (0.2196, 0.2255)	0
F14	Iterative filtering	0.4966; (0.4961, 0.4970)	0.4978; (0.4962, 0.4993)	0.4983; (0.4977, 0.4989)	0.4977; (0.4964, 0.4991)	0.4979; (0.4963, 0.4994)	0.4981; (0.4972, 0.4990)	0.0001
	EMD	0.4986; (0.4982, 0.4990)	0.4987; (0.4972, 0.5002)	0.4987; (0.4980, 0.4993)	0.4985; (0.4970, 0.4999)	0.4975; (0.4961, 0.4989)	0.4985; (0.4976, 0.4994)	0.4897
F15	Iterative filtering	0.5020; (0.5016, 0.5025)	0.5018; (0.5002, 0.5034)	0.5013; (0.5007, 0.5019)	0.5018; (0.5004, 0.5032)	0.5018; (0.5003, 0.5033)	0.5015; (0.5006, 0.5024)	0.0980
	EMD	0.5008; (0.5004, 0.5012)	0.5005; (0.4990, 0.5020)	0.5005; (0.4999, 0.5012)	0.5001; (0.4986, 0.5015)	0.5017; (0.5003, 0.5031)	0.5004; (0.4995, 0.5013)	0.2758
F16	Iterative filtering	1.8531; (1.8492, 1.8570)	1.5721; (1.5605, 1.5836)	1.3607; (1.3545, 1.3669)	1.2649; (1.2512, 1.2787)	1.1825; (1.1719, 1.1930)	1.5183; (1.5097, 1.5270)	0
	EMD	1.7500; (1.7477, 1.7523)	1.6225; (1.6157, 1.6294)	1.5202; (1.5159, 1.5245)	1.5356; (1.5248, 1.5465)	1.5244; (1.5178, 1.5311)	1.5962; (1.5902, 1.6022)	0
F17	Iterative filtering	0.0542; (0.0540, 0.0544)	0.0604; (0.0596, 0.0613)	0.0726; (0.0722, 0.0729)	0.0755; (0.0748, 0.0762)	0.0764; (0.0758, 0.0770)	0.0626; (0.0622, 0.0630)	0
	EMD	0.1355 (0.1350, 0.1361)	0.1425; (0.1398, 0.1451)	0.1919; (0.1907, 0.1931)	0.2048; (0.2022, 0.2075)	0.2048; (0.2028, 0.2069)	0.1617; (0.1596, 0.1638)	0
F18	Iterative filtering	1.8456; (1.8415, 1.8497)	1.5445; (1.5324, 1.5567)	1.3211; (1.3143, 1.3279)	1.2168; (1.2015, 1.2320)	1.1258; (1.1141, 1.1376)	1.4891; (1.4798, 1.4983)	0
	EMD	1.7916; (1.7889, 1.7942)	1.6355; (1.6279, 1.6431)	1.5356; (1.5305, 1.5407)	1.5616; (1.5496, 1.5736)	1.5658; (1.5589, 1.5726)	1.6119; (1.6054, 1.6184)	0

Table 3 continued

Feature	Method	Awake	S1	S2	S3	S4	REM	<i>p</i> value
F19	Iterative filtering	3.7361; (3.7192, 3.7530)	3.0263; (2.9936, 3.0590)	3.0776; (3.0620, 3.0932)	3.1724; (3.1319, 3.2128)	3.3016; (3.2588, 3.3443)	3.0680; (3.0466, 3.0893)	0
	EMD	2.7473; (2.7408, 2.7538)	2.5020; (2.4817, 2.5223)	2.1208; (2.1124, 2.1291)	2.1242; (2.1038, 2.1446)	2.1190; (2.0993, 2.1386)	2.4102; (2.3980, 2.4225)	0
F20	Iterative filtering	-0.0677; (-0.0761, -0.0592)	0.4675; (0.4495, 0.4855)	0.5842; (0.5741, 0.5942)	0.6812; (0.6584, 0.7040)	0.7903; (0.7710, 0.8096)	0.4708; (0.4550, 0.4866)	0
	EMD	-0.4358; (-0.4405, -0.4310)	-0.1517; (-0.1688, -0.1347)	-0.0439; (-0.0515, -0.0363)	-0.1098; (-0.1278, -0.0919)	-0.1507; (-0.1631, -0.1382)	-0.1695; (-0.1798, -0.1591)	0

work, the *p* values obtained using Kruskal–Wallis statistical test for six different groups (sleep stages) are reported in Table 3. It can be observed that the features obtained using EMD have resulted in relatively higher *p* values. Table 3 depicts the *p* values obtained from features computed from the first mode and first IMF.

2.5 Classification

The features extracted from EEG signals of different sleep stages are given as the input for different classifiers, and their performances are compared in terms of different performance parameters. The classifiers employed in this work are briefly discussed in the following sections.

2.5.1 Naïve Bayes

Naïve Bayes [31] is frequently used classifier that has a straightforward approach based on the application of Bayes' theorem. It provides simpler approach based on the probabilistic knowledge in order to predict test instances accurately. This algorithm assumes that the predictive attributes are conditionally independent and there are no hidden attributes which can affect the prediction process [31].

2.5.2 *k*-nearest neighbor

The *k*-nearest neighbor is one of the simplest and widely used classifier. It is one of the most straightforward approaches among different instance-based learning algorithms [69]. In this algorithm, similarity function like Euclidian distance is used to compute the similarity between training instances and the instances in classification record [7]. A record is maintained in order to store the classification performance and similarity results. In order to classify an instance, similarity with *k*-nearest neighbors are computed and the class corresponding to the maximum number of votes is assigned as output class of the instance.

2.5.3 C4.5 decision tree classifier

Among the tree-based classification algorithms, the C4.5 is most widely used inductive inference tool [17, 57]. The tree construction follows the top-down approach in which tree construction starts from a training set or tuples [60]. A tuple is the collection of attribute values and a class value. An attribute may have continuous or discrete values; however, class can have only discrete values. A decision tree consists of decision node and leafs. At each decision node, an attribute is specified which is tested for its ability

to classify a training sample [17]. Initially, the root node is associated with the whole training set, and the weight value for each case is set to 1. In order to construct a decision tree, the C4.5 algorithm employs divide and conquer approach [17, 57]. The attribute with the highest information gain is selected for the test at a node. Then, a child node is created for each possible outcome of the class. This process is repeated for each attribute associated with each node to select the best attribute for the node.

2.5.4 Multilayer perceptron

The multilayer perceptron is a particular type of neural network-based classifiers [45, 46]. This classifier employs a multilayer feed-forward neural network with one or more layers of nodes between the input and output layers. These nodes at different layers are interconnected through the weighted networks. Using different training algorithms, the parameters (weights) of the networks are optimized. The backpropagation algorithm [45] is commonly used for the purpose of parameter optimization. It uses the gradient search technique for minimizing a cost function. In this algorithm, initially, small random weights are selected. These weights are subsequently adjusted using class information during training until weights converge and cost function reduces to desirable limit [45].

2.5.5 Random forest

Random forest classifier [12] performs classification based on the collective decisions made by different classification trees. Each decision tree individually makes a decision about the class, and in order to make the final decision, a weight is assigned to each tree. Finally, to determine overall classification output, the class decision from every tree is considered. In order to build a tree, the random tree method is employed [18]. In the random forest algorithm, a random vector δ_n is assigned to the n th tree. For the generated vector δ_n , the distribution remains same as of previous random vectors but the random vector is produced independently of previous random vectors. Based on the training input data x and δ_n , a decision tree is grown which results in a tree classifier $H(x, \delta_n)$ [12]. The class is determined on the basis of margin function denoted by MG, which can be defined for a training set, randomly drawn from random vector distribution Y, X as [12]:

$$MG(X, Y) = av_n I[H_n(X) = Y] - \max_{j \neq Y} av_n I[H_n(X) = j] \tag{8}$$

where $H_n(X) = H(x, \delta_n)$ and $I(\cdot)$ denotes the indicator function [12]. The operator av_n indicates the average value. The larger value of margin is an indicator of more

confidence in the classification. The generalization error (GE) can be given by [12]:

$$GE = P_{X,Y}[MG(X, Y) < 0] \tag{9}$$

where $P_{X,Y}$ indicates probability over X, Y space. Strength and correlation are two parameters used to measure the accuracy of individual classifier and dependency between them. In this work, we have considered the set of the number of trees as {10, 20, 50, 80, 100, 120, 150, 180, 200, 250, 280, 300, 350} to perform classification and an optimum number of trees are decided for each case of multi-class classification. We have employed the Waikato environment for knowledge analysis (WEKA) [21] software implementation of naïve Bayes, k -nearest neighbor, C4.5 decision tree, multilayer perceptron, and random forest classifier for classification of sleep stages using EEG signals.

3 Results

In order to show the efficacy of the proposed methodology, it is studied on the EEG signals obtained from Sleep-EDF database. The number of epochs used for each class of data in the experiment is presented in Table 1. In the experimental study, the different sets of classes are used to formulate the different multi-class classification problems as listed in Table 4. The performance of the proposed methodology is evaluated in each multi-class classification case. Classification performance measures of these five different cases of multi-class classification problems are compared with the other existing methods.

In the Sleep-EDF dataset, PSG contains EEG signals recorded from Fpz–Cz and Pz–Oz channels. The EEG signals recorded from Pz–Oz channel are studied in this work. The suggested features are computed from first seven modes obtained from iterative filtering. We have performed different experiments for the studied multi-class classification problems designed from the available dataset. The sleep stages classification performance of the classifier is evaluated and presented in terms of accuracy (Acc), Sen, specificity (Spe), and Kappa measures [8, 16].

Table 4 The formulation of different cases of multi-class classification problems using different combinations of sleep stages

Multi-class classification problems	Relevant sleep stages
Two-class	W, Sleep (S1–S4, REM)
Three-class	W, NREM (S1–S4), REM
Four-class	W, (S1, S2), (S3, S4), REM
Five-class	W, S1, S2, (S3, S4), REM
Six-class	W, S1, S2, S3, S4, REM

The Kruskal–Wallis statistical test is applied to reduce the number of features used for the classification purpose. The features resulting $p < 0.05$ are selected for the classification purpose. In Table 3, p values are reported considering only six groups corresponding to the six classes of sleep stages. Similarly, p values are also obtained for five classes, four classes, three classes, and two classes. Then, for each case p values are tested for the significance of the features, and the features with $p > 0.05$ are removed from the feature set used for the classification purpose.

It is interesting to know that how many modes are suitable for feature extraction process. For this purpose, we start our experiment with the number of trees as 100 and employ different sets corresponding to varying number of modes for features extraction. The features extracted from different sets of modes are analyzed for their classification performance in each case of the multi-class classification problem. The set of modes which gives the best performance for 100 trees is used for the further experiment in each case. The classification accuracy for each case is given in Table 5. In Table 5, the q th mode extracted using iterative filtering is denoted by M_q . For example, the first mode is represented by M_1 . In a similar way, a set of modes is denoted by $M_x - M_y$. For example, a set of mode 1 to mode 7 is denoted by the $M_1 - M_7$. Then, the best classification accuracy is represented by bold entries in Table 5 corresponding to each multi-class classification case. The set of modes corresponding to the bold entries is used for

further experiments. To show the effectiveness of the iterative filtering over the EMD method, the classification accuracies obtained for the different combination of IMFs are also shown in Table 6. It can be clearly observed that the classification accuracy values are less than those obtained in the case of iterative filtering.

It is observed that the classification performance may vary for the different number of trees used in the random forest classifier. In order to select the appropriate number of trees, the classification is performed with varying number of trees. Therefore, for each case of multi-class classification, the number of trees used for classification is optimized based on the classification performance. The variation in the classification accuracy with the different number of trees is depicted in Figs. 5, 6, 7, 8 and 9. In Fig. 5, variation in the classification accuracy with the different number of trees is presented for the case of the six-class classification problem. In this case, highest classification accuracy is 90.02% when the number of trees used is 300. It can be observed from Fig. 6 that for the five-class classification problem, the maximum classification accuracy is 91.29% and the corresponding number of trees is 250. Similarly, for four- and three-class classification problems, the maximum classification accuracies obtained are 92.23 and 94.65%, respectively, and the corresponding number of trees are 300 and 180, respectively, as can be seen in Figs. 7 and 8. In the same way, the classification accuracy versus the number of trees plot, which is shown in

Table 5 Classification accuracy (%) for features extracted from different sets of modes extracted using iterative filtering for various cases of multi-class classification problem

Modes used for features extraction	Number of sleep stages for classification				
	Six stages	Five stages	Four stages	Three stages	Two stages
$M_1 - M_2$	88.68	89.98	90.73	94.08	97.58
$M_1 - M_3$	89.49	90.86	91.80	94.64	97.97
$M_1 - M_4$	89.75	90.03	92.07	94.45	97.93
$M_1 - M_5$	89.91	91.15	92.08	94.54	97.90
$M_1 - M_6$	89.55	90.92	91.97	94.41	97.87
$M_1 - M_7$	89.55	90.84	91.95	94.34	97.93

The number of trees considered is 100 in random forest classifier

Table 6 Classification accuracy (%) for features extracted from different sets of IMFs extracted using EMD method for various cases of multi-class classification problem

Modes used for features extraction	Number of sleep stages for classification				
	Six stages	Five stages	Four stages	Three stages	Two stages
$IMF_1 - IMF_2$	85.46	86.77	87.77	92.46	97.27
$IMF_1 - IMF_3$	87.11	88.62	89.43	93.74	97.40
$IMF_1 - IMF_4$	88.25	89.81	90.64	93.76	97.81
$IMF_1 - IMF_5$	88.61	90.16	90.91	93.71	97.88
$IMF_1 - IMF_6$	88.39	89.97	91.19	93.51	97.56
$IMF_1 - IMF_7$	88.35	90.05	90.62	93.54	97.54

The number of trees considered is 100 in random forest classifier

Fig. 5 Variation of classification accuracy as a function of number of trees for six-class classification problem in random forest classifier

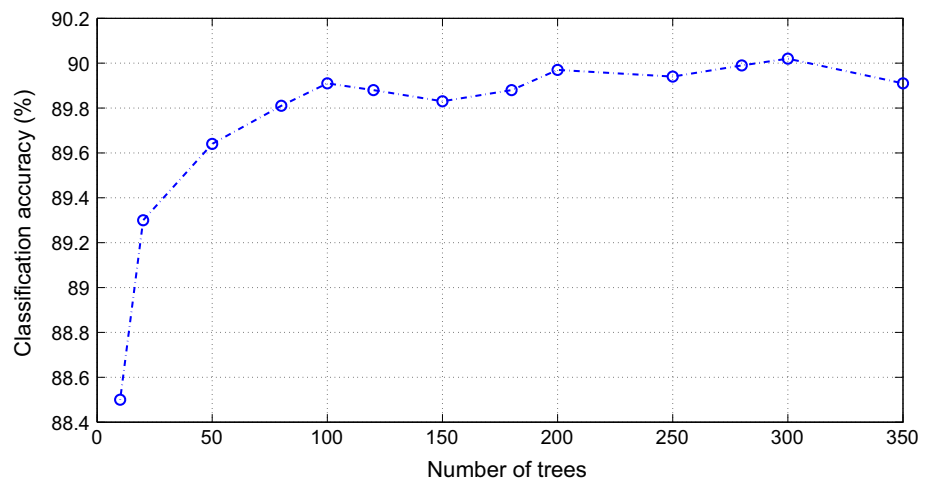


Fig. 6 Variation of classification accuracy as a function of number of trees for five-class classification case in random forest classifier

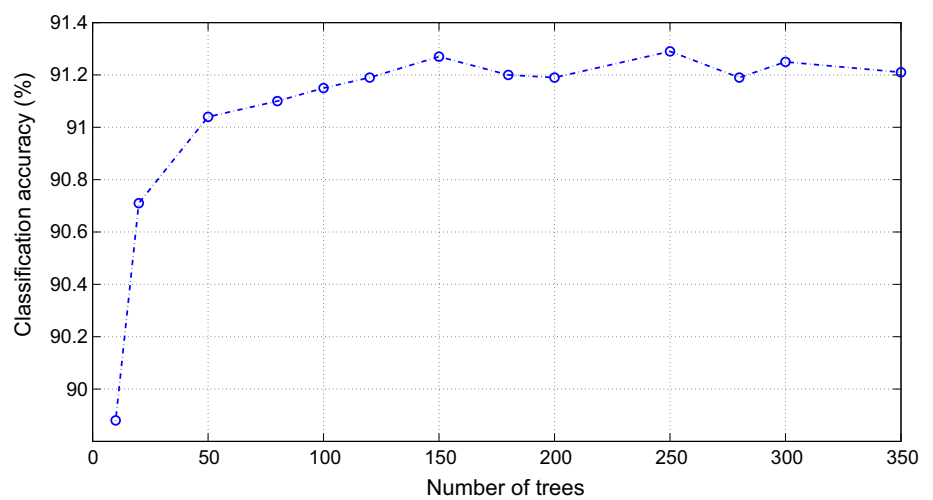


Fig. 7 Variation of classification accuracy as a function of number of trees for four-class classification problem in random forest classifier

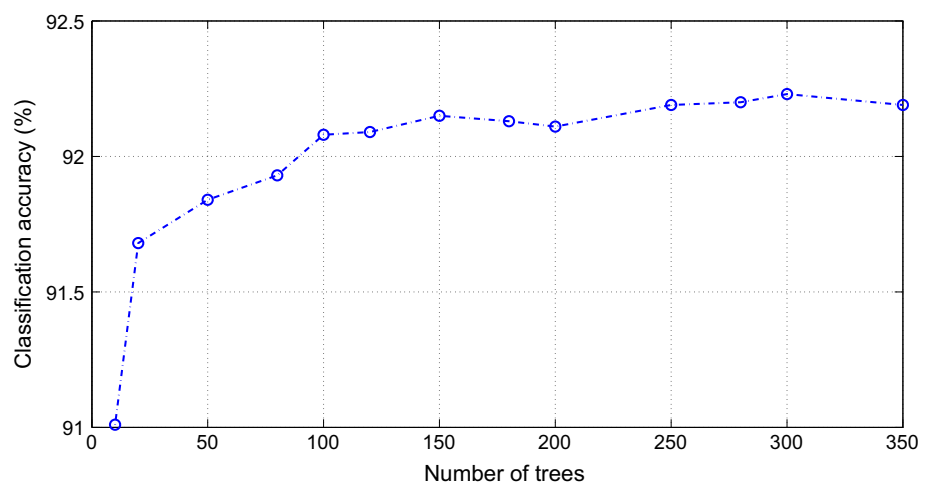


Fig. 9, indicates that the maximum classification accuracy is 98.02% for 250 trees.

The confusion matrix of five-class classification problem is presented in Table 7. It provides the Sen and Spe of

each individual class epochs. Similarly, confusion matrix for the six-class classification is shown in Table 8. It can be observed that from Tables 7 and 8 that the Sen values of the REM sleep stage detection are significantly higher,

Fig. 8 Variation of classification accuracy as a function of number of trees for three-class classification case in random forest classifier

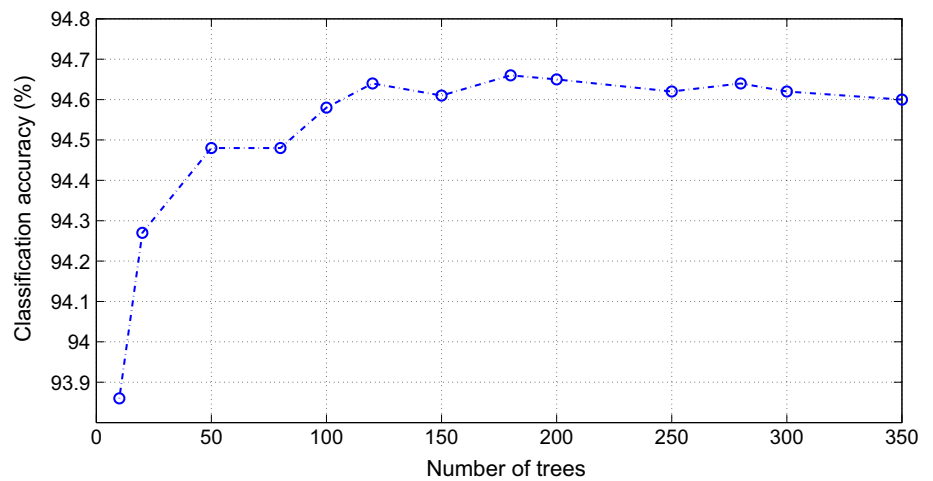


Fig. 9 Variation of classification accuracy as a function of number of trees for two-class classification problem in random forest classifier

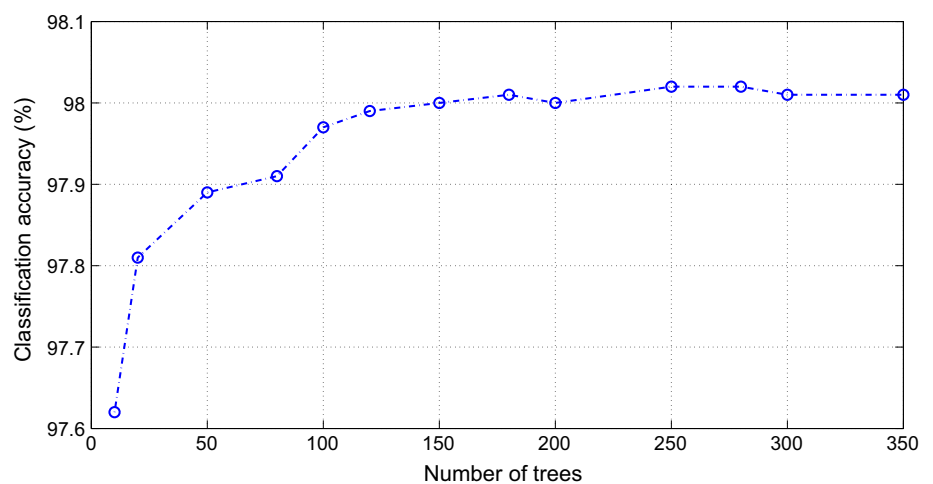


Table 7 Confusion matrix with Sen and Spe for five-class classification case

	Automatic classification					Sen (%)	Spe (%)
	W	S1	S2	S3–S4	REM		
<i>Expert annotation</i>							
W	7944	11	12	6	30	99.30	95.90
S1	183	113	123	4	181	18.7	99.80
S2	48	4	3334	149	86	92.10	95.30
S3–S4	13	0	198	1088	0	83.80	98.9
REM	52	11	207	0	1339	83.2	97.80

Table 8 Confusion matrix with Sen and Spe for six-class classification case

	Automatic classification						Sen (%)	Spe (%)
	W	S1	S2	S3	S4	REM		
<i>Expert annotation</i>								
W	7942	13	14	1	6	27	99.2	95.9
S1	179	114	124	3	1	183	18.9	99.80
S2	50	6	3350	101	30	84	92.5	95.20
S3	11	0	209	352	100	0	52.4	98.70
S4	3	0	12	90	522	0	83.3	99.10
REM	52	13	198	0	0	1346	83.70	97.80

which shows good detection ability of the classifier for REM sleep stage.

Tables 7 and 8 show that most of the S1 stage epochs are recognized as the REM sleep stage. In other words, they are harder to distinguish. In order to investigate the performance of the proposed methodology, we further trained the classifier to distinguish S1 from REM sleep

stages. Table 9 lists the obtained results and comparison of results with the method proposed by Zhu et al. [73]. It is found that when trained with proposed features, the random forest classifier can discriminate between S1 and REM sleep stages with an accuracy of 84.00%. This accuracy is obtained when classifier trained with features

Table 9 Comparison of the results obtained for classification between S1 and REM sleep stages

Authors	Classification accuracy (%)	Kappa coefficient
Zhu et al. [73]	78.8	0.34
Proposed method	84.00	0.54

extracted from first seven modes and 200 trees are considered for classification. The achieved classification accuracy is better than that of reported by Zhu et al. [73] for classification between S1 and REM stages. They obtained the value of Kappa coefficient as 0.34 for classification of S1 and REM stages. In the case of proposed methodology, the obtained value of Kappa coefficient is 0.54, which is better than that of achieved in Zhu et al. [73].

The performance of the proposed features is investigated for different classifiers. The classification accuracy obtained from different classifiers is shown in Table 10. The compared classifiers include the naïve Bayes, *k*-nearest neighbor, multilayer perceptron, C4.5 decision tree, and random forest. The purpose of the comparison is to investigate the suitability of different classifiers for studied features. All the experiments are performed using WEKA software for multi-class classification problems obtained from different sleep stages. The naïve Bayes is the simplest classifier but resulted in poor classification accuracy as compared to other classifiers. While experimenting with the *k*-nearest neighbor classifier, the value of *k* is varied from 1 to 20. The best accuracy achieved for each multi-class classification problem is mentioned in Table 10. The multilayer perceptron is a neural network-based classifier.

Table 10 Comparison of various classifiers for studied features

Classifier	Two-class	Three-class	Four-class	Five-class	Six-class
Naïve Bayes	89.01	75.46	75.29	74.18	71.80
<i>k</i> -nearest neighbor	95.91	90.49	87.13	85.10	83.56
Multilayer perceptron	96.84	93.04	89.12	88.72	87.13
C4.5	96.67	91.69	88.04	87.18	84.85
Random forest	98.02	94.60	92.16	91.13	89.74

Table 11 Comparison of results reported in other existing methods with proposed method

Authors	Multi-class classification					Epochs used for training/testing
	Two-class	Three-class	Four-class	Five-class	Six-class	
Ronzhina et al. [59]	98.62	90.31	81.55	–	76.70	Tenfold cross-validation
Zhu et al. [73]	97.90	92.60	89.30	88.90	87.50	7481/7482
Hassan and Bhuiyan [23]	99.48	94.10	92.14	90.69	86.89	7592/7596
Hassan and Bhuiyan [22]	97.50	94.80	92.11	91.50	90.38	7594/7594
Proposed methodology	98.02	94.66	92.29	91.13	90.02	Tenfold cross-validation

In the case of multilayer perceptron classifier, we performed experiments with the different number of hidden layers. Starting with one hidden layer, it is observed that on increasing the number of hidden layers, the classification performance improves; however, at the same time, time consumption gets increased. Therefore, we chose ten hidden layers for experiments. In the case of the C4.5 classifier, with default parameters its speed is better than that of random forest classifier, however, it resulted in lower classification accuracy than that of random forest classifier. Finally, we obtained best classification performance with random forest classifier as can be observed from Table 10.

The performance of the proposed methodology is compared with the other recent state-of-art sleep scoring techniques reported in the literature. The methods presented in these studies are evaluated on the Sleep-EDF database. Table 11 provides comparison of the classification accuracies achieved in different existing methods and the proposed method. In Table 12, the Kappa values obtained using proposed methodology are compared for each case of multi-class classification with results of the methodology presented by Zhu et al. [73].

4 Discussion

In this work, a technique for classification of sleep stages by employing nonstationary signal analysis method is developed based on single channel EEG signal. This method can be useful for computer-aided sleep stage detection. The DESA is used to capture the variation in the

Table 12 The Kappa statistics corresponding to highest accuracy obtained in different multi-class classification cases

Multi-class classification case	Kappa statistics	
	Zhu et al. [73]	Proposed methodology
Two-class	0.96	0.96
Three-class	0.87	0.9066
Four-class	0.83	0.8745
Five-class	0.83	0.8619
Six-class	0.81	0.8429

amplitude and frequency of the modes obtained using iterative filtering. The statistical parameters are employed for identifying characteristics of the different sleep stages from amplitude envelope and instantaneous frequency functions of the modes. The significance of the features is examined using Kruskal–Wallis statistical test. The random forest classifier is employed with tenfold cross-validation procedure. The selection of the suitable number of modes in order to extract the features is addressed. The procedure for selection of the appropriate number of decision trees is given. In the following paragraphs, we will discuss some of the important attributes of the developed algorithm.

The EEG signal is nonlinear and nonstationary in nature [2, 49, 52]; therefore, in this work, the nonstationary signal decomposition technique namely iterative filtering is used to decompose the sleep EEG epochs. The main difference between the iterative filtering and EMD is that the iterative filtering involves low pass filtering to extract modes, whereas the EMD method uses spline interpolation approach for obtaining components of the signals. The use of low pass filtering in iterative filtering for determining modes is a simpler in principle and makes it more suitable for real-time implementation as compared to EMD method. In addition, iterative filtering is suitable for horizontal comparison between modes of the data. The term horizontal comparison means that different signals belonging to the same class when decomposed in the same number of modes, the modes of the same index contain comparable information [67]. In other words, the modes of the same index belonging to different signals can be compared in terms of particular characteristic feature. Moreover, iterative filtering is less sensitive for the small perturbation as compared to EMD due to use of low pass filters at each stage [67]. In EMD method, the small perturbation may result in the different set of IMFs [71]. However, in order to further compare iterative filtering over EMD, we have also implemented the proposed algorithm by replacing iterative filtering with EMD method. Then, the different multi-class classification cases with random forest classifier are compared. The presented

results show the better performance of iterative filtering-based classification technique. These results support the suitability of the iterative filtering in the proposed method.

It has been observed that different sleep stages can be characterized by variation in frequency content and amplitude [20]. For example, the S1 sleep stage can be characterized by low voltage and the mixed frequency with the highest amplitude in frequency range 2–7 Hz. Similarly, the sleep stage S3 can have 2 Hz or slower oscillations with the amplitude of 75 mV. The motivation for the application of the DESA is to capture these variations in frequency and amplitude by decomposing the mode in amplitude envelope and instantaneous frequency functions. The obtained amplitude envelope and instantaneous frequency functions of the mode can represent the variations in frequency and amplitude for different sleep stages. Further, these variations are quantified using various features. The TEO can be defined for three consecutive samples of a mode. This fact also makes iterative filtering algorithm more suitable for decomposition because it can better handle effects of perturbation. Therefore, the amplitude envelope and instantaneous frequency functions can be estimated more accurately.

To analyze the sleep physiology and pathophysiology, the gold standard is the PSG recorded in the laboratory [35]. The sleep scoring using PSG is basically performed in hospital settings or clinical laboratory environment. At such places, the subject has to wait for a considerable period of time due to long waiting list [41]. Moreover, the sleep recording may require the subject to stay overnight in the hospital. It may affect the sleep quality of the subject due to the unfamiliar environment in the hospital or sleep clinic [6, 41]. In addition, the sleep scoring at laboratory may be expensive, require expertise, and it is tedious and time-consuming process [1]. Therefore, there is an increasing interest for the home-based sleep assessment systems [35]. There are several commercially available software system for the sleep monitoring purpose. These sleep monitoring systems make the decision based on different physiological signals. The example of such sleep monitoring system is *Zeo* device [35]. This system consists of the elastic headband which has sensors for multiple channel recordings such as EEG, EMG, and EOG signals. Therefore, the system developed based on single EEG channel can provide a convenient solution for home-based sleep scoring and analysis with relative ease.

The methods developed using the multiple physiological signals such as ECG, EEG, EMG present some difficult challenges. The subject preparation requires a long-time and complicated procedure for recording of multiple signals. For example, electrodes often need to be glued for proper recording of the ECG signals. It may also include

the abrasion of subject's skin area for low impedance [62]. It often poses hindrance and inconvenience for the common activities such as body movements. However, this work employs only single channel EEG recordings for the identification of sleep stages and therefore, overcome these limitations up to a large extent. Also during recordings of multiple physiological signals, ECG, EOG, and EMG can introduce artifacts in EEG signals [54]. This may require preprocessing and artifact rejection using filtering procedure or correlation analysis. Therefore, the use of single channel EEG signal can reduce the significant time for artifacts and noise reduction. The use of single channel-based algorithm is an important advantage of the proposed methodology because it needs less power consumption, hence, makes it suitable for implementation as power efficient device. It can also facilitate portable, wearable, and feasible implementation of the device for sleep quality evaluation [1].

The detection of REM is important for the analysis of REM sleep behavior disorder [30]. The deprivation of the REM can be associated with other neurodegenerative disorders like Parkinson's disease and dementia [30]. The characteristics of REM sleep stage is low voltage and mixed frequency in EEG [20]. In PSG recordings, the REM stage can be identified with the aid of EOG and chin EMG signals [62]. The rapid eye movements that occur during REM stage records in EOG. Similarly, the level of chin muscles decreases during REM stage and can be observed in chin EMG. However, it is found difficult to distinguish REM from W and S1 sleep stages using automatic sleep scoring [10, 54].

In Zhu et al. [73], the confusion matrix for five-class and six-class classification cases shows Sen values for REM sleep stage detection as 76.2 and 76.2%, respectively. Similarly in Hassan and Bhuiyan [23], the confusion matrix for five-class and six-class classification problem shows Sen values for REM sleep stage which are 80.80 and 78.39%, respectively. In another study [22], reported the Sen of REM sleep stage detection as 82.11 and 83.60% for five-class and six-class classification problems, respectively. It is evident from Tables 7 and 8 that the Sen of the REM sleep stage detection obtained using proposed method is better than former above-mentioned two studies and equivalent to results of Hassan and Bhuiyan [22]. However, it should be noted that all three studies use nearly 50% data for training and remaining data for testing the classifier. On the other hand, in the proposed study, the tenfold cross-validation procedure is followed to obtain classification performance. The Sen values for S1 detection are better in the case of the methodologies presented in [22, 23, 73]. The reason for the less Sen of the S1 detection can be attributed to the less number of epochs as compared to awake and REM stages. The less Sen of S1

detection is also evident from the study presented in Ronzhina et al. [59]. Therefore, the classification is also performed separately for S1 and REM sleep stages. The results obtained for the classification of S1 and REM epochs also indicate the importance of proposed methodology in distinguishing the S1 and REM epochs.

The comparison of classification performance of different methods, in terms of classification accuracy, is presented in Table 11. It can be observed that the classification accuracies for six-class to three-class classification problems obtained using proposed method are better than that of the first three methods presented in Table 11. In these multi-class classification problems, the performance of the proposed method is equivalent to the methods presented in Hassan and Bhuiyan [22]. Ronzhina et al. [59] employed the cross-validation procedure; however, in other methods training and testing data are split differently as mentioned in Table 11. Instead, the proposed method employed tenfold cross-validation performance. For two-class classification, results are slightly better in the case of the study presented in Hassan and Bhuiyan [23] and Ronzhina et al. [59]. We have used the same recordings as studied by Zhu et al. [73] and Hassan and Bhuiyan [22], and the proposed methodology shows better performance than their methodology in terms of tenfold cross-validation classification accuracy. In the proposed methodology, to compute the classification performance, the classification has been performed using tenfold cross-validation with WEKA software. The interobserver bias in the classification can be addressed based on the Kappa values [42]. In Landis and Koch [42], the authors discussed the importance of the Kappa values with respect to the strength of agreement among observers on the classification of the individual class. The value of $\text{Kappa} \geq 0.81$ is considered as an almost perfect agreement among different observers. Kappa values, listed in Table 12, suggest the almost perfect performance of the proposed methodology for each multi-class classification problem. We have also performed the comparison of proposed methodology in terms of Kappa coefficient with those presented in Zhu et al. [73]. The Kappa values show better performance of the proposed methodology as the Kappa values are higher for last four multi-class classification cases and in the two-class classification problem, the Kappa values are equal in case of both studies.

In this work, the EEG signals recorded from the eight subjects are considered for all the experiments. The EEG epochs for a sleep stage taken from different subjects are combined to be treated as a single class. The effect of subject variability over different sleep stages is not analyzed in this work. In fact, the developed method is aimed to detect the sleep stages without having any prior information about the patient. Also, the approach of sleep stage

detection adopted in the proposed methodology is consistent with other studies carried out on the same database.

In future, a larger number of patients can be included for further study of proposed method. The sleep constitutes the different duration of various sleep stages. The duration of the sleep also varies from subjects to subjects; therefore, the number of epochs of various sleep stages varies too. It is obvious from Table 1 that the number of epochs for different sleep stages is distributed unevenly. For example, the number of epochs in S1, S3, and S4 is very less as compared to epochs belong to awake. This leads to the class imbalance problem in the sleep EEG classification. The problem of class imbalance can be handled by resampling techniques such as synthetic minority oversampling technique (SMOTE) [14]. It should be noted that the performance of the proposed method is significant as compared to the other compared state-of-art methods. It is difficult to obtain perfect classification for all the studied multiclass classification cases, may be due to class imbalance problem.

5 Conclusion

In this paper, iterative filtering-based decomposition is studied for automated classification of sleep stages using EEG signals. The TEO is used to extract the features from the modes obtained by applying the iterative filtering technique. In order to extract amplitude envelope and instantaneous frequency functions, the DESA is employed which uses the TEO. It is found that the features extracted from amplitude envelope and instantaneous frequency functions of the modes obtained using iterative filtering can efficiently discriminate the various sleep stages from EEG signals. The classification is performed using the five different classifiers namely, naïve Bayes, k -nearest neighbor, multilayer perceptron, C4.5 decision tree, and random forest classifiers. The performance of random forest classifier is found most suitable for classification of sleep stages. We also presented a strategy to find optimal choice of modes to extract features for each case of multi-class classification. The performance of the classifier is tested for the various number of decision trees, and suitable choice for the number of decision trees is reported for each case of the multi-class classification problem. The results show that the proposed methodology provides improved and more robust performance. The results were also compared with other existing methodologies for two to six-class classification problems. We have also shown the effectiveness of features for classification between the S1 and REM sleep stages which is found better than the previous study. The proposed study only uses single channel EEG signal-based strategy for sleep stages classification. The strategy based on single channel EEG signals has several advantages as compared to the strategy based on

multichannel or multiple physiological signals. The multichannel signal-based methodology may increase the complexity of the algorithm and makes it complicated to implement in wearable devices for automatic sleep quality assessment. On the other hand, the multiple physiological signal-based method may suffer through the complicated process of subject preparation for signal acquisition. It also requires many number of electrodes to be placed which makes it unsuitable for a hassle-free process of recordings. We conclude that with the obtained results and mentioned advantages, the proposed methodology can be suitable for designing automated sleep scoring devices.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

References

1. Aboalayon KAI, Faezipour M, Almuhammadi WS, Moslehpour S (2016) Sleep stage classification using EEG signal analysis: a comprehensive survey and new investigation. *Entropy* 18(9):1–31
2. Acharya UR, Faust O, Kannathal N, Chua T, Laxminarayan S (2005) Non-linear analysis of EEG signals at various sleep stages. *Comput Methods Programs Biomed* 80(1):37–45
3. Acharya UR, Bhat S, Faust O, Adeli H, Chua ECP, Lim WJE, Koh JEW (2015) Nonlinear dynamics measures for automated EEG-based sleep stage detection. *Eur Neurol* 74:268–287
4. Adeli H, Ghosh-Dastidar S, Dadmehr N (2007) A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy. *IEEE Trans Biomed Eng* 54(2):205–211
5. Agarwal R, Gotman J (2001) Computer-assisted sleep staging. *IEEE Trans Biomed Eng* 48(12):1412–1423
6. Agnew HW, Webb WB, Williams RL (1966) The first night effect: an EEG study of sleep. *Psychophysiology* 2(3):263–266
7. Aha D, Kibler D (1991) Instance-based learning algorithms. *Mach Learn* 6:37–66
8. Azar AT, El-Said SA (2014) Performance analysis of support vector machines classifiers in breast cancer mammography recognition. *Neural Comput Appl* 24(5):1163–1177
9. Bajaj V, Pachori RB (2013) Automatic classification of sleep stages based on the time-frequency image of EEG signals. *Comput Methods Programs Biomed* 112(3):320–328
10. Berthomier C, Drouot X, Herman-Stoca M, Berthomier P, Prado J, Bokar-Thire D, Benoit O, Mattout J, d'Ortho MP (2007) Automatic analysis of single-channel sleep EEG: Validation in healthy individuals. *Sleep* 30(11):1587–1595
11. Bouchikhi A, Boudraa AO (2012) Multicomponent AM-FM signals analysis based on EMD-B-splines ESA. *Sig Process* 92(9):2214–2228
12. Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
13. Brennan M, Palaniswami M, Kamen P (2001) Do existing measures of Poincaré plot geometry reflect nonlinear features of heart rate variability? *IEEE Trans Biomed Eng* 48(11):1342–1347
14. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 16:321–357
15. Cicone A, Liu J, Zhou H (2016) Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis. *Appl Comput Harmonic Anal* 41(2):384–411

16. Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20:37–46
17. Farid DM, Zhang L, Rahman CM, Hossain M, Strachan R (2014) Hybrid decision tree and naïve Bayes classifiers for multi-class classification tasks. *Expert Syst Appl* 41(4, Part 2):1937–1946
18. 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. *Comput Methods Programs Biomed* 108(1):10–19
19. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* 101(23):e215–e220
20. Güneş S, Polat K, Yosunkaya Şebnem (2010) Efficient sleep stage recognition system based on EEG signal using *k*-means clustering based feature weighting. *Expert Syst Appl* 37(12):7922–7928
21. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The weka data mining software: an update. *SIGKDD Explor* 11(1):10–18
22. Hassan AR, Bhuiyan MIH (2016a) A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features. *J Neurosci Methods* 271(15):107–118
23. Hassan AR, Bhuiyan MIH (2016b) Computer-aided sleep staging using complete ensemble empirical mode decomposition with adaptive noise and bootstrap aggregating. *Biomed Signal Process Control* 24:1–10
24. Held CM, Heiss JE, Estevez PA, Perez CA, Garrido M, Algarin C, Peirano P (2006) Extracting fuzzy rules from polysomnographic recordings for infant sleep classification. *IEEE Trans Biomed Eng* 53(10):1954–1962
25. Hsu YL, Yang YT, Wang JS, Hsu CY (2013) Automatic sleep stage recurrent neural classifier using energy features of EEG signals. *Neurocomputing* 104:105–114
26. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen NC, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond A Math Phys Eng Sci* 454(1971):903–995
27. Iber C, Ancoli-Israel S, Chesson A, Quan S (eds) (2007) *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. American Academy of Sleep Medicine, Westchester
28. Imtiaz SA, Rodriguez-Villegas E (2014) A low computational cost algorithm for REM sleep detection using single channel EEG. *Ann Biomed Eng* 42(11):2344–2359
29. Iranzo A, Santamaría J, Tolosa E (2009) The clinical and pathophysiological relevance of REM sleep behavior disorder in neurodegenerative diseases. *Sleep Med Rev* 13(6):385–401
30. Iranzo A, Molinuevo JL, Santamaría J, Serradell M, Martí MJ, Valldeoriola F, Tolosa E (2006b) Rapid-eye-movement sleep behaviour disorder as an early marker for a neurodegenerative disorder: a descriptive study. *Lancet Neurol* 5(7):572–577
31. John GH, Langley P (1995) Estimating continuous distributions in Bayesian classifiers. In: Eleventh conference on uncertainty in artificial intelligence. Morgan Kaufmann, San Mateo, pp 338–345
32. Kaiser JF (1990) On a simple algorithm to calculate the ‘energy’ of a signal. *Int Conf Acoust Speech Signal Process* 1:381–384
33. Kaleem MF, Sugavaneswaran L, Guergachi A, Krishnan S (2010) Application of empirical mode decomposition and Teager energy operator to EEG signals for mental task classification. In: Annual international conference of the IEEE engineering in medicine and biology, pp 4590–4593
34. Kayikcioglu T, Maleki M, Eroglu K (2015) Fast and accurate PLS-based classification of EEG sleep using single channel data. *Expert Syst Appl* 42(21):7825–7830
35. Kelly JM, Strecker RE, Bianchi MT (2012) Recent developments in home sleep-monitoring devices. *International Scholarly Research Notices*, pp 1–10, article ID 768794
36. Kemp B, Zwinderman AH, Tuk B, Kamphuisen HAC, Oberyé JLL (2000) Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Trans Biomed Eng* 47(9):1185–1194
37. Kirkwood BR, Sterne JA (2003) *Essential medical statistics*, 2nd edn. Blackwell, Oxford
38. Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the 14th international joint conference on artificial intelligence, vol 2, pp 1137–1145
39. Krakovská A, Mezeiová K (2011) Automatic sleep scoring: a search for an optimal combination of measures. *Artif Intell Med* 53(1):25–33
40. Kruskal WH, Wallis WA (1952) Use of ranks in one-criterion variance analysis. *J Am Stat Assoc* 47(260):583–621
41. Lan KC, Chang DW, Kuo CE, Wei MZ, Li YH, Shaw FZ, Liang SF (2015) Using off-the-shelf lossy compression for wireless home sleep staging. *J Neurosci Methods* 246:142–152
42. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33(1):159–174
43. Liang SF, Kuo CE, Hu YH, Pan YH, Wang YH (2012) Automatic stage scoring of single-channel sleep EEG by using multiscale entropy and autoregressive models. *IEEE Trans Instrum Meas* 61(6):1649–1657
44. Lin L, Wang Y, Zhou H (2009) Iterative filtering as an alternative algorithm for empirical mode decomposition. *Adv Adapt Data Anal* 01(04):543–560
45. Lippmann R (1987) An introduction to computing with neural nets. *IEEE ASSP Mag* 4(2):4–22
46. Madyastha RK, Aazhang B (1994) An algorithm for training multilayer perceptrons for data classification and function interpolation. *IEEE Trans Circuits Syst I Fundam Theory Appl* 41(12):866–875
47. Maragos P, Kaiser JF, Quatieri TF (1993) Energy separation in signal modulations with application to speech analysis. *IEEE Trans Signal Process* 41(10):3024–3051
48. Motamedi-Fakhr S, Moshrefi-Torbati M, Hill M, Hill CM, White PR (2014) Signal processing techniques applied to human sleep EEG signals: a review. *Biomed Signal Process Control* 10:21–33
49. Pachori RB (2008) Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition. *Res Lett Signal Proc* 2008:1–5
50. Pachori RB, Gangashetty SV (2010a) AM-FM model based approach for detection of glottal closure instants. In: 10th international conference on information science, signal processing and their applications, pp 266–269
51. Pachori RB, Gangashetty SV (2010b) Detection of voice onset time using FB expansion and AM-FM model. In: 10th international conference on information science, signal processing and their applications, pp 149–152
52. Pachori RB, Sircar P (2008) EEG signal analysis using FB expansion and second-order linear TVAR process. *Sig Process* 88(2):415–420
53. Pachori RB, Sircar P (2010) Analysis of multicomponent AM-FM signals using FB-DESA method. *Digit Signal Proc* 20(1):42–62
54. Penzel T, Conradt R (2000) Computer based sleep recording and analysis. *Sleep Med Rev* 4(2):131–148

55. Piskorski J, Guzik P (2007) Geometry of the Poincaré plot of RR intervals and its asymmetry in healthy adults. *Physiol Meas* 28(3):287–300
56. Potamianos A, Maragos P (1994) A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation. *Sig Process* 37(1):95–120
57. Quinlan JR (1986) Induction of decision trees. *Mach Learn* 1(1):81–106
58. Rechtschaffen A, Kales A (eds) (1968) A manual of standardized terminology. Public Health Service, U.S. Government Printing Office, Washington, DC, Techniques and Scoring System for Sleep Stages of Human Subjects
59. Ronzhina M, Janoušek O, Kolářová J, Nováková M, Honzík P, Provazník I (2012) Sleep scoring using artificial neural networks. *Sleep Med Rev* 16(3):251–263
60. Ruggieri S (2002) Efficient C4.5. *IEEE Trans Knowl Data Eng* 14(2):438–444
61. Sharmila V, Krishna EH, Reddy KA (2013) Cumulant based Teager energy operator for ECG signal modeling. In: International conference on advances in computing, communications and informatics, pp 1959–1963
62. Sleep Technology: Technical Guideline (2012) Standard polysomnography. Technical report, American Association of Sleep Technologies. <https://www.aastweb.org/technical-guide-lines>. Accessed 09 Jan 2017
63. Spurrier JD (2003) On the null distribution of the Kruskal-Wallis statistic. *J Nonparametr Stat* 15(6):685–691
64. Tsinalis O, Matthews PM, Guo Y (2016) Automatic sleep stage scoring using time-frequency analysis and stacked sparse autoencoders. *Ann Biomed Eng* 44(5):1587–1597
65. Villanueva J, Shaffer DR, Philip J, Chaparro CA, Erdjument-Bromage H, Olshen AB, Fleisher M, Lilja H, Brogi E, Boyd J, Sanchez-Carbayo M, Holland EC, Cordon-Cardo C, Scher HI, Tempst P (2006) Differential exoprotease activities confer tumor-specific serum peptidome patterns. *J Clin Investig* 116(1):271–284
66. Šušmáková K, Krakovská A (2008) Discrimination ability of individual measures used in sleep stages classification. *Artif Intell Med* 44(3):261–277
67. Wang Y, Wei GW, Yang S (2011) Iterative filtering decomposition based on local spectral evolution kernel. *J Sci Comput* 50(3):629–664
68. Wang Y, Wei GW, Yang S (2012) Mode decomposition evolution equations. *J Sci Comput* 50(3):495–518
69. Wilson DR, Martinez TR (2000) Reduction techniques for instance-based learning algorithms. *Mach Learn* 38(3):257–286
70. Wu HT, Talmon R, Lo YL (2015) Assess sleep stage by modern signal processing techniques. *IEEE Trans Biomed Eng* 62(4):1159–1168
71. Wu Z, Huang NE (2009) Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv Adapt Data Anal* 1(1):1–41
72. Zhou G, Hansen JHL, Kaiser JF (2001) Nonlinear feature based classification of speech under stress. *IEEE Trans Speech Audio Process* 9(3):201–216
73. Zhu G, Li Y, Wen P (2014) Analysis and classification of sleep stages based on difference visibility graph from a single-channel EEG signal. *IEEE J Biomed Health Inform* 18(6):1813–1821