Classification Manner Utilizing Electroencephalography Signals to Investigate Waveforms

Wessam Al-Salman

Ali Basim Al-Khafaji

Mishall Al-Zubaidie

Classification Manner Utilizing Electroencephalography Signals to Investigate Waveforms

Abstract

Waveform detection has been an area of continuous investigation for many years. One important waveform in sleep stage 2 is the *k*-complex. Numerous researchers have created various strategies for auto-*k*-complex detection; some of these strategies state that the automated detection techniques are adequate. Because of its analytically relevant resolution, the Electroencephalogram (EEG) is a commonly utilized technique to analyze the *k*-complexes in order to understand the nervous system activity of the brain. Several researchers have classified waveforms using EEGs in a variety of ways. It appears that the majority of the waveform detection had limitations. The necessary analyzes took a lot of time to complete, no execution time was indicated in any of the earlier research, and they were too complex for real-world use. Furthermore, it was revealed that numerous experiments were done without window size and were utilized to detect waveform characteristics. Additionally, the research used one or two evaluation instruments to analyze the performance outcomes. For the dataset, a maximum accuracy of 94% to 75% was reported. Because of its significance, several analysts have developed an automated technique to use EEG data to study *k*-complexes. This work proposes a novel approach to feature detection for *k*-complexes utilizing a least square support vector machine (LS-SVM) classifier. The sliding window method divides EEG signals into a number of segments. Subsequently, distinct feature sets are obtained from every time interval. Every EEG segment was visible in the obtained twenty-seven features as vectors. That means, twenty-seven features were extracted using the Katz algorithm and the Tunable Q-factor wavelet transform for each segment. These features were analyzed, to select the most important features, using an Analysis of Variance (ANOVA) and the F-test. Finally, the vector of features was used as input to the LS-SVM classifier. When it came to identifying events of (non) *k*-complexes, the suggested novel technique demonstrated noteworthy performance results with sensitivity, accuracy, and specificity of 98.3%, 96.5%, and 91.6%, respectively. This high accuracy rate has not been discovered in any method yet. When compared to other classifiers and methods in this field of research, the LS-SVM classifier approach yielded the best results.

Keywords

EEG; k-complexes; sleep; sliding windows; support-vector machine; Signals

RESEARCH PAPER

Classification Manner Utilizing Electroencephalography Signals to Investigate Waveforms

Wessam Al-Salman, Ali B. Al-Khafaji, Mishall Al-Zubaidie*

Department of Computer Sciences, Education College for Pure Sciences, University of Thi-Qar, Thi-Qar, 64001, Iraq

Abstract

Waveform detection has been an area of continuous investigation for many years. One important waveform in sleep stage 2 is the *k*-complex. Numerous researchers have created various strategies for auto-*k*-complex detection; some of these strategies state that the automated detection techniques are adequate. Because of its analytically relevant resolution, the Electroencephalogram (EEG) is a commonly utilized technique to analyze the *k*-complexes in order to understand the nervous system activity of the brain. Several researchers have classified waveforms using EEGs in a variety of ways. It appears that the majority of the waveform detection had limitations. The necessary analyzes took a lot of time to complete, no execution time was indicated in any of the earlier research, and they were too complex for real-world use. Furthermore, it was revealed that numerous experiments were done without window size and were utilized to detect waveform characteristics. Additionally, the research used one or two evaluation instruments to analyze the performance outcomes. For the dataset, a maximum accuracy of 94%−75% was reported. Because of its significance, several analysts have developed an automated technique to use EEG data to study k-complexes. This work proposes a novel approach to feature detection for k-complexes utilizing a least square support vector machine (LS-SVM) classifier. The sliding window method divides EEG signals into a number of segments. Subsequently, distinct feature sets are obtained from every time interval. Every EEG segment was visible in the obtained twenty-seven features as vectors. That means, twenty-seven features were extracted using the Katz algorithm and the Tunable Q-factor wavelet transform for each segment. These features were analyzed, to select the most important features, using an Analysis of Variance (ANOVA) and the F-test. Finally, the vector of features was used as input to the LS-SVM classifier. When it came to identifying events of (non) *k*-complexes, the suggested novel technique demonstrated noteworthy performance results with sensitivity, accuracy, and specificity of 98.3%, 96.5%, and 91.6%, respectively. This high accuracy rate has not been discovered in any method yet. When compared to other classifiers and methods in this field of research, the LS-SVM classifier approach yielded the best results.

Keywords: EEG, *k*-complexes, Sleep, Sliding windows, Support-vector machine, Signals

1. Introduction

The brain's ability to heal and restore people's physical and mental health is thought to depend on the stage of sleep. This aids in the body's ability to heal itself and boosts the immune system. When people's jobs and personal lives are impacted by how they sleep, it serves an important purpose in life [1]. Brain disarray can cause long-lasting issues that can harm people's ability to function physically and mentally. 50−70 m people were discovered to suffer from sleep disorders in 2003, including chronic sleeplessness and sleep apnoea, according to the US Health Organization [2,3]. A sleep specialist would typically stage sleep visually throughout the night using the Rechtschaffen & Kales function or the American Academy of Sleep Medicine (AASM) features [4−6]. The examination

Received 12 July 2024; revised 4 October 2024; accepted 6 October 2024.
Available online 4 November 2024

* Corresponding author.
E-mail address: mishall_zubaidie@utq.edu.iq (M. Al-Zubaidie).

Peer review under responsibility of University of Kerbala.

https://doi.org/10.33640/2405-609X.3377

of the sleep state is tiring, error-prone, and expensive. The stages of sleep were distinguished as rapid eye movement (REM) and non-rapid eye movement (NREM) [7–10]. S1-Stage 1, S2-Stage 2, S3-Stage 3, S4-Stage 4 make up NREM [11,12]. According to numerous studies, the EEG signals show distinct characteristics and waveforms at each stage of sleep that show whether a person is awake or asleep right now [7]. These sleep stages' waveforms often depict any alterations to the muscles and any movement of the brain's neurons [3,4]. Fig. 1 provides an illustration of the characteristics of these sleep wave levels. The existence of *k*-complexes in S2 in signals from electroencephalograms is a crucial feature. The waves are described as a prominent series of sharp negative waves followed by strong positive waves. The *k*-complexes are reported to last between 0.5 and 1.5 s (s), with a maximal amplitude of over 75 μV [13,14]. According to other studies, the minimum value of [15–17] is 100 μV, and the maximum and minimum k-complex durations are 1 s and 3 s, respectively.

The gold standard for *k*-complexes identification is generally thought to be sleep expert visually graded methods. Although visual recording for *k*-complexes was once thought to be the conventional approach, it has a number of drawbacks and restrictions. A specific one is biased, challenging, tiring, expensive, and inaccurate. Nevertheless, since conclusions reached by the two sleep experts may differ, even while recording the same sleep, visually determining the morphology of the recordings of *k*-complexes in EEG can be potentially erroneous and untrustworthy [15]. As a result, increasing numbers of researchers are concentrating on creating a spontaneous *k*-complex finding

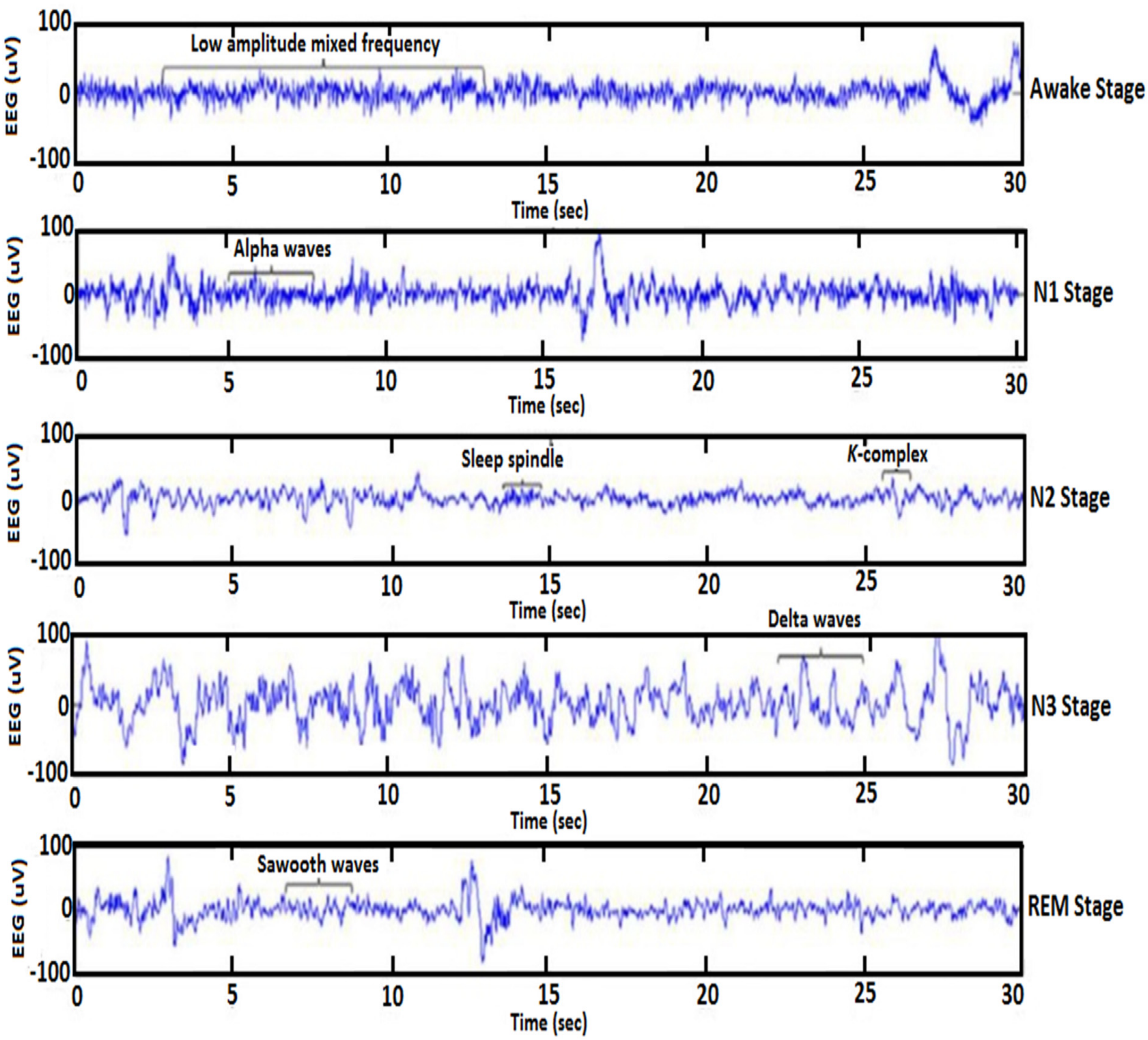


Fig. 1. Wave features through different sleep stages in signals of EEG.

approach to fasten treatment and lessen the neurologist's workload. Additionally, a trustworthy approach to figuring out *k*-S2 complexes would save time and provide an unbiased assessment of the sleep stage. In this field of study, many domains of time-frequency or time-domain, signal-based, single-channel, multichannel physiological, and multiple approaches have been presented to identify automatic sleep *k*-complexes.

The key contributions and novelty of the proposed approach are given below:

- The principal contribution of this work is the development of a novel waveform detection algorithm based on fractal dimension and the Tunable Q-Factor Wavelet Transform (TQWT) to extract significant features from EEG signals used to identify the key elements of sleep stage 2 from EEG data. This method is used as a platform to help physicians and neurologists identify sleep problems automatically using EEG records. Experts find it useful to make clinical judgments as well. It was created using a distinct set of features and combined with the LS-SVM classifier. The FD method combines the signals, and it is used to extract association characteristics from EEG data without any prior information. Single channel signals are the focus of our method's performance analysis. This is just another crucial component of our approach. The implemented method offers significant accuracy for waveform recognition, reporting a correct detection average of 98.3%. Thus, our approach contributes to the improvement of the systematic approach to automatically score sleep phases, helping physicians and neurologists identify and treat sleep problems. Moreover, our method's primary advantage relates to the feature extraction process. With this approach, a vast amount of EEG data may be minimized into small datasets by selecting the most representative data across all dataset segments while accounting for observation variability. This study may operate with large amounts of data while requiring little resources and computations due to data reduction.
- Using a common dataset, the effectiveness of our model was tested, and when compared to other models, the database yielded the best and most accurate findings. The performance of detection is greatly developed for all waveforms investigation using various sets of characteristics.
- The suggested technique is fast since it relies on selection characteristics based on ANOVA and F-tests, which need minimum computation time. Our suggested method used EEG data to characterize signals that are characteristic for the identification of waveforms in sleep stages and sleep-related works.
- This study uses a publicly available EEG database so that our findings may be expanded upon, examined, and contrasted with those of other researchers.
- We have achieved high classification accuracy in 6-fold cross-validation using our approach, which is also applicable in one offline database. In comparison to other techniques, this is the best categorization accuracy yet documented in this discipline.
- The performance of our technique was evaluated using a variety of classification measuring instruments, such as F-score, cross-validation, ROC curve, specificity, accuracy, sensitivity, and kappa coefficient. These evaluation tools double-checked the viability of techniques to identify sleep waveform characteristics in EEG recordings. Thus, EOGs, ECGs, and bio-signals of epileptic seizures may all be tested and classified using our technique.
- In conclusion, an effective approach for extracting features from the entire EEG database was presented, utilizing Katz's algorithm and the Q-Factor Wavelet Transform to identify waveforms that enhance both system and classification performance.

The study roadmap is described in this article: In Section 1, a thorough introduction is given. Works related to this topic are included in Section 2. We provide a description of experimental data in Section 3. Section 4 outlines our proposed method process. Section 5 studies feature extraction and signal data. Section 6 investigates statistical analysis for our proposed method. Classification algorithms are illustrated in Section 7. Section 8 presents experimental findings with details on the discussion of results and comparisons. The limitations of the proposed solution are included in Section 9. Finally, the study's findings and projected trends are presented in Section 10.

2. Related work

Some methods are used to discover signals of EEG *k*-complexes. Al-Hadeethi et al. [18], for instance, used a matching approach to filter data in order to determine the morphology of the *k*-complexes in raw EEG data. An adjustable Q-factor wavelet transform (TQWT) was proposed by Lajnef et al. [19] to study *k*-complexes. Bankman et al. [16] focused on using an informal neural network classifier having 14 features to discover the characteristics of S2 if EEG sleep signal. The average of the sensitivity results obtained was 90%. A neural network was also applied by Jansen and Desai [13] to discover *k*-complex and non-*k*-complex fragments from EEG signals. Many researchers investigated *k*-complexes in EEG sleep by applying various methods such as an artificial neural network [13,16], and nonlinear characteristics concerning fractal dimensions, Fuzzy thresholds [1], and pattern-matched wavelet methods [20], time-frequency and deep neural [21], multitaper method [22], and deep learning [23], clustering method and a classifier of neural network [24]. Stranded in 68%−92% of cases, the accuracy of finding *k*-complexes is medium. Therefore, the largest hurdle in *k*-complexes detection is now effectively classifying imbalanced data. The suggested model also has a promising performance when handling severely unbalanced data. Medical practitioners such as neurologists may find it helpful in the identification and management of sleep disorders. To circumvent this limitation, the purpose of this work is to suggest new methods for the investigation of EEG *k*-complex signaling.

In this work, an effort is made to create an effective technique for recognizing EEG signals that have *k*-complexes of different features coupled using an LS-SVM. In this study, the window size was used to divide every signal to some number of epochs utilizing the method of sliding windows. Then, other groups of characteristics, including frequency, fractal, and statistical time aspects, were extracted from every segment. To display the *k*-complexes sections and those non-*k*-complexes, the classifier of LS-SVM was fed all the excerpted features after analysis. To evaluate the effectiveness of the suggested strategy, the LS-SVM's results were compared to those of other various classifiers. Then, it was compared to earlier research. Our results demonstrated that specific feature sets might classify the *k*-complexes.

3. Experimental data description

The EEG dream database is used by the researchers as a baseline for EEG data sets when doing the studies. Researchers can access the data set and other information for free at http://www.tcts.fpms.ac.be/devuyst/Database/DatabaseSpindles [9]. The effectiveness of our approach was assessed using the Sleep-EDF database. Additionally, this data set was collected from the Dream database that MONS University made available to the TCTS lab. It is also accessible through Bruxelles University [1] from the sleep laboratory at de-Charleroi. Four men and six women were EEG signal recorders. They were between the ages of 20 and 47 (see Fig. 2): 30-min EEG signals from the EEG central channel were recorded, these signals are (C3- A1) or (Cz-

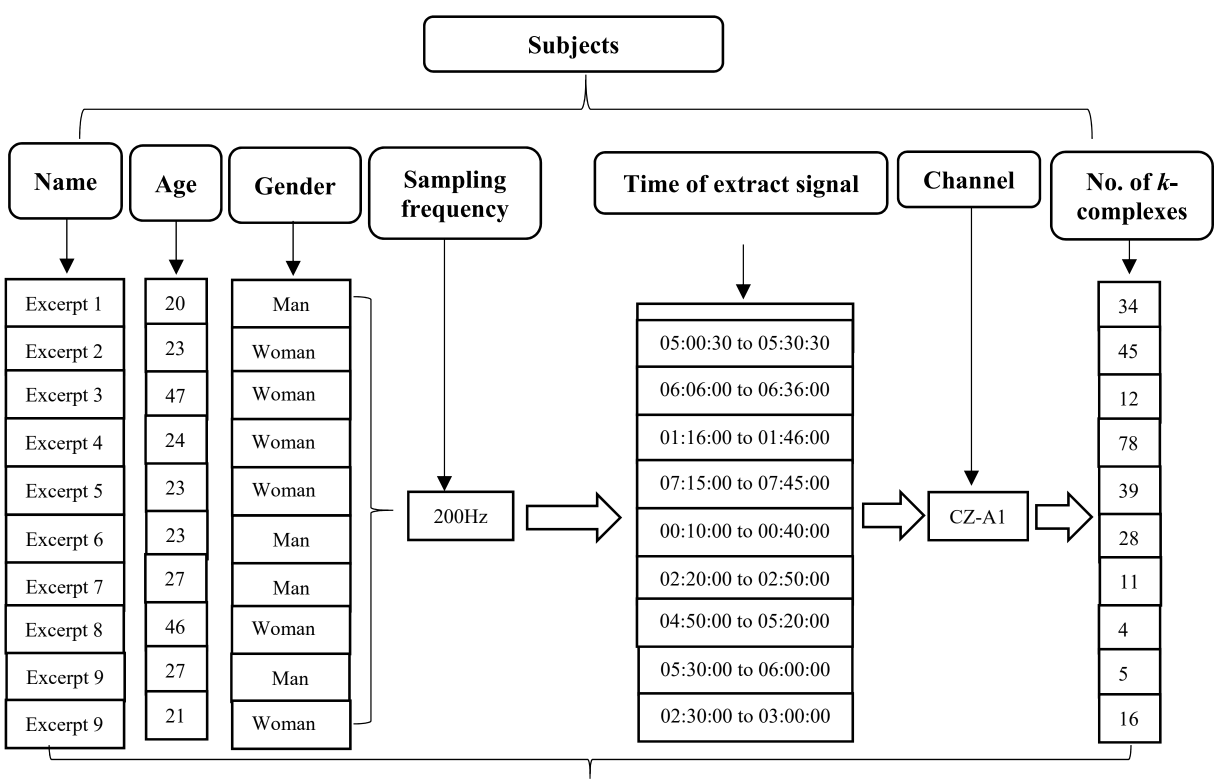


Fig. 2. Details of k-complexes database.

A1). Signals are sampled at a rate of 200 Hz. Every EEG recording comprises one EMG submental channel in addition to 3 channels of EEG: channel 1 is C3- A1 or channel CZ- A1, channel 2 is O1- A1, and channel 3 is FP1- A1. Two sleep specialists visually identified the EEG data collected through the study of signals of EEG.

1st sleep expert took 10 recordings, 2nd sleep expert rated the 5−10 dataset. A duration of 1 s was constantly *k*-complexes providing. This paper uses the 1st 10 recordings because they provide a more accurate measurement of the duration. Fig. 3 displays an illustration of a stage two sleep EEG with *k*-complexes and without them.

4. Proposed method

In this effort, a useful technique to recognize EEG signal *k*-complexes was provided. As seen in Fig. 4, every EEG signal in part 1 was divided into manageable parts using the sliding window technique. Li et al. [25] divided the epileptic EEG data using this method. In their research, Zaidi, and Farooq [3] used the method of sliding windows to identify the features of S2 in EEG signals. Ashokkumar et al. [26] utilized the sliding window method in order to assess the anesthesia degree in the EEG. Zaidi, and Farooq [3] looked at several sizes of sliding windows to identify and discover any sleep spindles in the EEG signals. According to the outcomes of the simulation, the sliding window size in this study was experimentally determined through training phases. After thorough testing, the size of the window was managed at 0.5 s and 0.4 s of overlap. Various sections of characteristics have been retrieved by applying a dimension fraction technique and transformation of wavelet with a tunable Q-factor in Section 2, which is the features extraction stage. A crucial step to developing the functionality and reducing algorithm complexity is to lower the feature dimensionality.

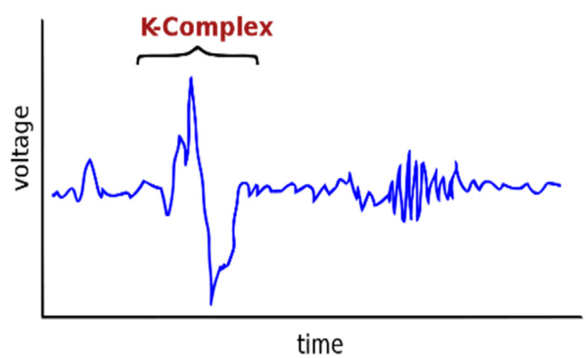


Fig. 3. Example of EEG signals of k-complexes.

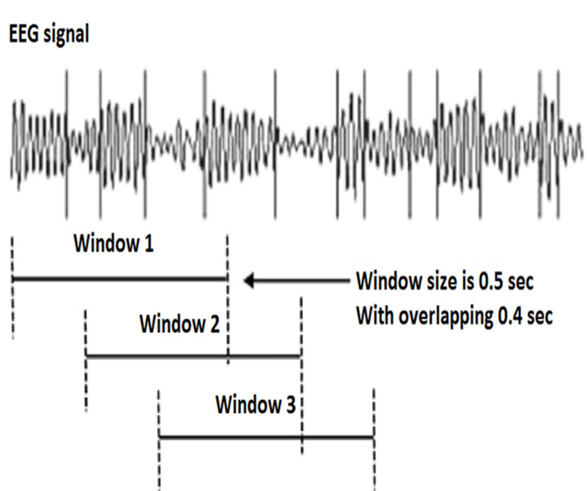


Fig. 4. Example of EEG waves being divided into 0.5-second-long pieces using the sliding window technique.

The fundamental merits were then employed to characterize *k*-complexes by applying the SVM after the retrieved features had been assessed using computational hypothesis testing to show each segment of EEG for 0.5 s. We sent the retrieved k means features, Bayes who is a nerves network classifier tests the functioning of our suggested technique. These classifiers were chosen because of their effectiveness. We performed a 6-cross comparison and validation of various approaches in this area of study in order to manage additional evaluation. The outcomes showed that our suggested method is capable of bringing out various aspects of EEG signals. Additionally, it has the ability to identify the essential components of S2, such as the high accuracy level of *k*-complexes. These findings support the sleeping issue diagnosis by neurologists and sleep specialists. The three steps of our method's structure are shown in Fig. 5.

4.1. Features extraction

The feature extraction technique aims to minimize the large volume dimensionality of the signal data. This needs to be done without leaving out any important details. For instance, feature extraction is regarded as a crucial step to achieve crucial findings in the classification stage. Applying the fractal and tunable Q-Factor Wavelet Transform, various characteristics set "Twenty-three" are extracted from every 0.5 segments of EEG in this study. In some studies, waveforms such as a *k*-complex are defined as temporary transient waveforms that are observed by a negative sharp wave followed by a positive sharp wave with a duration between 0.5 s and 1.5 s. Some studies have reported that the maximum time duration of this waveform is between 1 s and 3 s

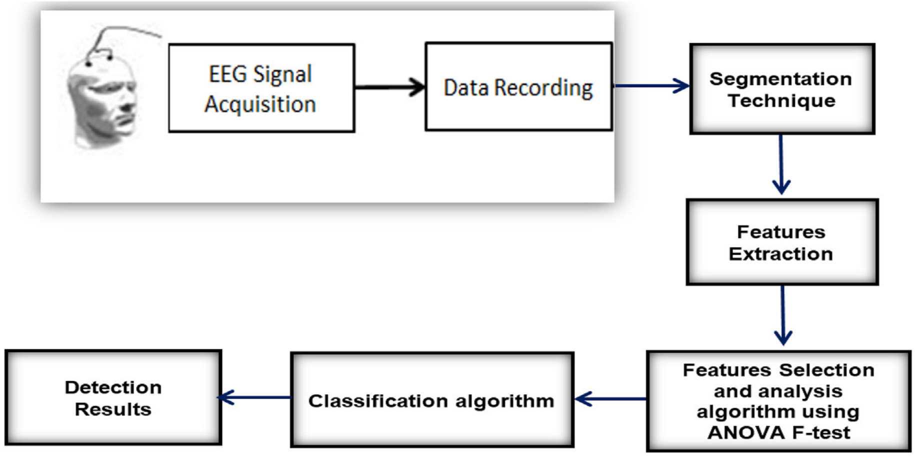


Fig. 5. A representation of the proposed to find k-complexes in the EEG data.

(Bankman et al., 1992; Cash et al., 2009; Devuyst et al., 2010; Jansen & Desai 1994). In addition, to detect all waveforms in sleep stage 2, different window sizes of 0.5 s, 1 s, 1.5 s, and 2 s were tested. During the training phase, our findings show that the 0.5 s achieved higher results compared with others.

Prior to feeding the retrieved features into the classifier, we evaluated the statistical hypothesis to evaluate the features statistically. A statistical hypothesis test is primarily used to identify and differentiate the potential of selected traits and to indicate whether or not these statistically differentiate the potential. Al-Kabir et al.'s [27] previously published work from 2023 states that selecting the best characteristics is important for the EEG classification challenge. The results show that selecting characteristics accurately enhances the classification performance. Statistical hypothesis testing is hence a popular method for selecting characteristics. One-way analysis variance (ANOVA) was utilized in this work to assess the statistical significance of the features distinguishing potential and to ascertain if feature values in different classes varied substantially. It is usually used to compare the characteristics of different populations.

$$FD=\frac{Log10\left(\frac{L}{a}\right)}{Log10\left(\frac{L}{a}\right)+Log10\left(\frac{d}{L}\right)} \tag{1}$$

ANOVA scores the "f-test" variance between classes and variance within the class when examining features. The degree of class division is displayed by the "f-ratio." Musthafa et al. (2024) [28] have provided further information on ANOVA. Every experimental test was carried out using the statistics toolbox in Matlab. Consequently, if $p < 0.05$, a difference is considered statistically significant, because of this, features with p-values larger than 0.05 are deemed to be non-significant and are removed from the feature matrices.

4.1.1. Fractal dimension

By using Katz's approach to calculate the dimension fraction of each EEG segment, fractal dimension characteristics (FD) were recovered. It is used to calculate the ratio between the maximum distance d and the curve length L. This is done at every position along the curve, from start to finish. L is derived by adding together all of the Euclidean distance between subsequent places. This is how the signals' fractal dimension is calculated:

Where

- a refers to the uniform distance between all following places.
- L means the total length curve or the total of the Euclidean distance between two consecutive points.
- d stands for the diameter of curves.
- L and d are illustrated by:

$$L=\sum_{j=1}^{n}\left[\left(x_{j+1}-x\right)^2+\left(y_{j+1}-y\right)^2\right]^{1/2} \tag{2}$$

$$d=\max\left(\left[\left(x_{j+1}-x\right)^2+\left(y_{j+1}-y\right)^2\right]^{1/2}\right);j=1,2,3\ldots,n \tag{3}$$

In this work, the FDs for each epoch were calculated using Katz's technique. The value of the FDs is predicted to be a 1.0 to 2.0 fraction.

4.1.2. Tunable Q-factor wavelet transform (TQWT)

In this study, TQWT was employed to identify the characteristics of EEG signals. The TQWT often depends on various factors affecting the conclusions [8,29,30]. TQWT, which is reported by Selesnick (2011), is a transform of discrete wavelet (DWT) that is adaptable. The TQWT uses a bank of 2-channel filters, similar to the DWT, that contains a low-pass with parameter α and a high-pass filter with parameter, employing programmable *Q*-factors, to separate the signal of EEG into sustained components. The low-pass filter's sustained component output is taken into consideration as signal input for the following 2-channel filter bank for additional analysis. The signals output are the drifting components of the filter of high-pass for each output layer. Fig. 6 shows a straightforward instance of a transformation wavelet with a J level. TQWT is covered in greater detail in [14,31−33]. These conditions are:

1. *Q*-factor: The band-pass filter's width is determined by this parameter. The TQWT reaction enables flexibility by managing and modifying this wavelet transform value. The sustained components extraction is more efficient the higher the *Q*-factor. In the meantime, the reacted waveform on the basis of a lower *Q*-factor is useful to gain the drifting component's properties.
2. Number of decomposition levels (*J*): Input signals are divided into sub-bands of *J*+1 if *J* is the filter band number. *J* sub-bands for each of these bands were obtained through their highpass filters, and one came from the final level filter band's low-pass filter. The time domain waveform widens and the characteristics sharply increase as the decomposition level is raised.
3. Sampling rate or redundancy (*R*).

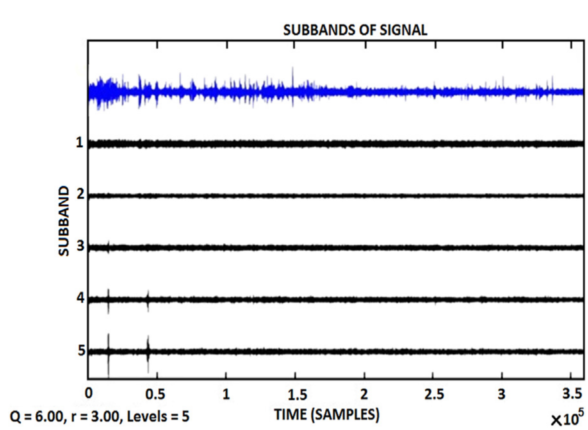


Fig. 6. One example of wavelet transformation with Q-factor and J level.

4. *R* and *Q* can be computed based on low pass scaling and high pass scaling factors (β and σ).

$$R = \frac{\beta}{\sigma - 1} \tag{4}$$

$$Q = \frac{\beta - 2}{\beta} \tag{5}$$

5. TQWT was used to divide EEG signals into many levels or sub-bands (SBs) applying the parameters of input (*Q*, *R*, and *J*). A sampling frequency and a (L_PS) low pass sub-band and (H_PS) high pass sub-band σfs and βfs, respectively, as shown in Fig. 7.
6. A filter of low pass h_1 (ω scaling σ is applied to make the L_PS, and the H_PS is made from a filter of high pass h_0 (ω) with scaling β. Then, H_PS is decayed into L_PS and H_PS as higher and lower sub-bands at *j*th level, respectively. Equations (6) and (7) can be used to mathematically express them. The low-pass filter's sustained component output is taken into consideration as the signal input for the following 2-channel filter bank for further analysis. The output signals are the drifting groups that the filter of high-pass for every layer outputs. Fig. 5 shows a straightforward instance of a transformation wavelet with *J* level.

Low pass filter $h0(\omega)$

$$= \begin{cases} 1, & if\ |\omega| \le (1-\beta)\pi, \\ \Theta \dfrac{\omega + (\beta-1)\pi}{\sigma+\beta-1}, & if\ (1-\beta)\pi < |\omega| < \alpha\pi, \\ 0, & if\ \sigma\pi \le |\omega| \le \pi \end{cases} \tag{6}$$

High pass filter $h1(\omega)$

$$= \begin{cases} 1, & if\ |\omega| \le (1-\beta)\pi, \\ \Theta\left(\dfrac{\alpha\pi - \omega}{\sigma+\beta-1}\right), & if\ (1-\beta)\pi < |\omega| < \alpha\pi, \\ 0, & if\ \sigma\pi \le |\omega| \le \pi \end{cases} \tag{7}$$

7. The mathematical formula for the frequency that corresponds with low and high pass filters in TQWT can be gained mathematically.

In this work, the *Q*-value was fixed at 6 after extensive testing, and it was discovered that SB3 offered the needed properties for locating the EEG *k*-complexes. Finally, various parts of characteristics

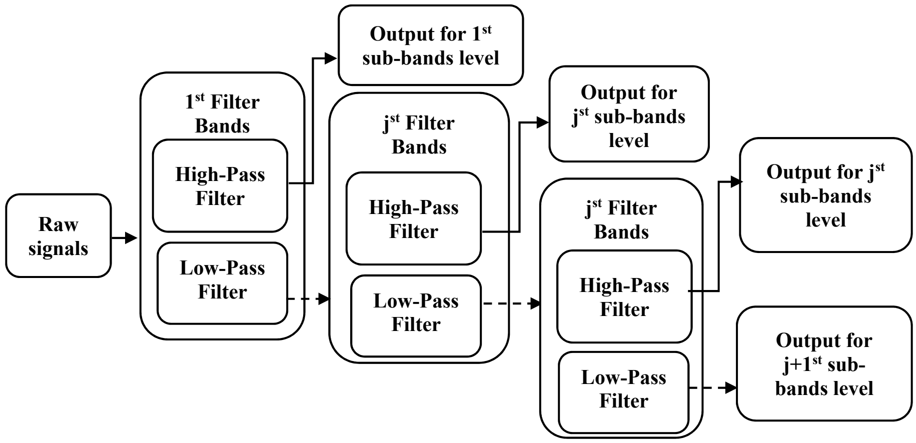


Fig. 7. EEG signal decomposes using TQWT into low pass subband and high pass subband at jth levels.

entropy, mean energy, maximum, Hurst exponent (H), minimum, median, mode, mean, range, first quartile, skewness, variation, second quartile, standard deviation, shape factor, kurtosis, crest factor, Zero-crossing rate, margin factor, impulse factor, short energy, and time centroid were taken out from SB3 for every segment and then combined with others fractal dimension and frequency features peak frequency, mean frequency, mean power, median frequency which are taken out using Katz's approach and power spectrum density, respectively. The PSD was defined as:

$$P(\omega) = \sum_{n=0}^{N=1} x(n) e^{-j\frac{2\pi}{N}\omega n} \omega = 0, 1, 2, \ldots\ N-1 \tag{8}$$

Where ω is the frequency of normalization angular and n refers to the time index. The peak frequency, median frequency, mean power, and mean frequency were derived from PSD as the most significant distinctive aspects. To identify k-complexes, those traits are combined with other features. As was already noted, one crucial step to lessen algorithm complexity and boost performance is to lower the size of EEG data. All of the collected features were further examined using P-value and then significant sections of characteristics were employed as input to the classifier LS-SVM to identify (non) k-complex segments, see Table 1. The proposed characteristics P-value, which were calculated from sub-bands of TQWT, demonstrate how k-complex features differ from non-k-complex features. The test is run with a 96% level of confidence. According to Table 1's results, the qualities that are bolded are significant ($p \leq 0.05$), as shown in Fig. 8, and a deviation is only considered significant in statistics $p \leq 0.05$. The findings demonstrate that distinct sets of features performed much better than other characteristics in classifying k-complexes. These findings lead us to the conclusion that not all of the features had successful k-complex detection capabilities. Therefore, in order to enhance k-complex detection performance and shorten processing time, some of these features must be chosen. Table 2 shows the parameters of the classifiers utilized in the proposed research experiments.

4.2. Statistical analysis

The functioning of our method was assessed utilizing a variety of statistical analyzes, including ROC

Table 1. P-Value of the proposed features based on ANOVA f-test.

No.	Features	F-ratio	P-value	Result
1	Maximum	26.6978958	0.4033e-06	S
2	Range	42.53727	1.136e-08	S
3	Standard deviation	13.94123	0.000396	S
4	Minimum	26.92904	2.19263e-06	S
5	Mean	1.39979	0.241482	NS
6	Short energy	10.22504	2.54e-16	S
7	Zero-crossing rate	8.52190	0.52e-10	S
8	Time centroid	11.03799	1.36e-08	S
9	Impulse factor	24.36407	0.016979	S
10	Margin factor	14.21629	9.47e-13	S
11	Shape factor	9.68142	0.001373	S
12	Crest factor	5.20659	0.683969	NS
13	Mean energy	9.56247	2.27e-18	S
14	Entropy	8.89128	0.0009451	S
15	Hurst exponent	6.51018	6.49e-34	S
16	Mode	0.00256	0.959829	NS
17	Median	0.01223	0.912279	NS
18	First quartile	0.03899	0.844079	NS
19	Fractal feature	6.49223	0.000803	S
20	Second quartile	1.83682	0.179944	NS
21	Variation	8.41826	0.005043	S
22	Skewness	2.15407	0.134247	NS
23	Kurtosis	19.8159	0.97e-4539	S
24	Mean frequency	9.96944	0.0024	S
25	Median frequency	2.89938	0.093319	NS
26	Peak frequency	13.24693	0.000537	S
27	Mean power	0.00034	0.985259	NS

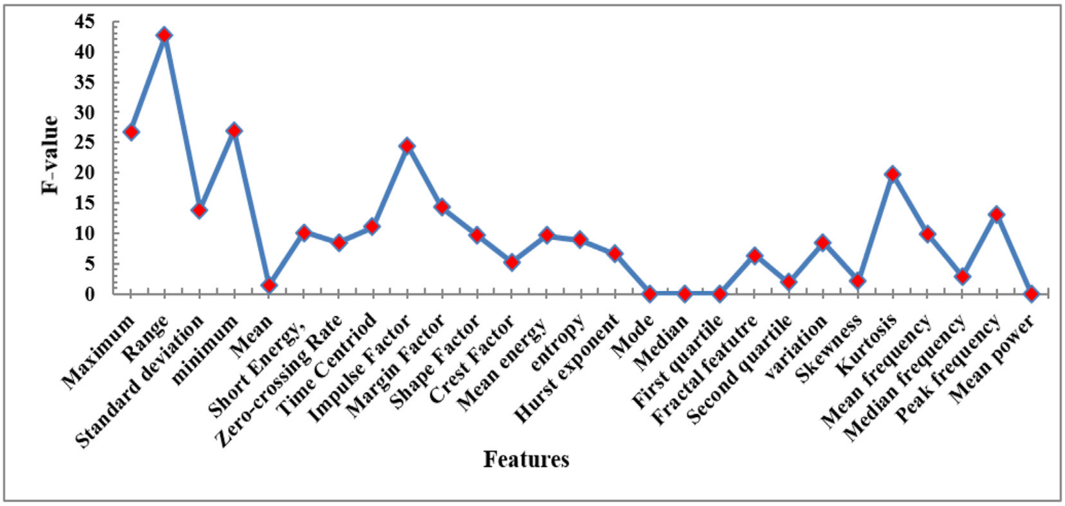


Fig. 8. The features ranking by using F-value; the high F-values (F-ratio> 2.9) refer to more efficient features.

Table 2. The parameters of the classifiers used in the experiments in this research.

Classifiers	Parameters and values
LS-SVM	The RBF kernel is used in this study. ($\gamma = 1$, $\sigma 2 = 10$),
k-means	k, c_i and x_k, where k is the number of clusters, $k = 2$, while c_i is the centre of clusters and $c_i = 1$, and x_k is the data points
Naïve bayes	Class nodes are represented by the EEG k-complexes, whilst the feature nodes stand for the frequency, statistic, and non-linear data.
ANN	The value of α is set to 1. The number of iterations is set to 1000, the target error is set to 10^{e-5}. The learning rate is set to 0 and activation function of $y\ (x) = 1/(1 + e^{-\ \alpha x})$.

(for additional information on the measures), sensitivity, cross-valuation, accuracy, and specificity, see [3,4,34]).

1. ***k*-cross validation**: It is used in this study to gauge and evaluate the categorization quality. It is divided into k-fold equal-size subgroups of a dataset. K represents the total amount of the inputs. Every fold of input is applied to the test, but others are used as training inputs and for testing (validation). This process is carried out n times.

$$\text{Performance} = \frac{1}{k} = \sum_{1}^{k} Accuracy^{k} \tag{9}$$

2. **Sensitivity**: It is a metric applied to quantitatively evaluate the classification effectiveness of findings utilizing the positive decision's actual number or cases.

$$\text{Sensitivity} = TP/(TP + FN) \tag{10}$$

3. **Specificity**: It is utilized to calculate the negative case rate by applying the decision number of true negative/negative cases number.

$$\text{Specificity} = TN/(TP + FP) \tag{11}$$

4. **Accuracy**: It is used to calculate the total number of accurate forecasts. It is used to evaluate a classifier's performance and accuracy. It specifies the proportion of true classified I stances to the final number of cases.

$$\text{Accuracy} = (TN + TP) = (TN + FN + FP + TP) \tag{12}$$

Where TP represents correctly detecting k-complexes in EEG signal, FN represents incorrectly identifying k-complexes, TN represents true negative detection number and FP represents true positive detection number.

4.3. Classification algorithms

The supervised learning method of machine learning has been used to categorize binary data using LS-SVM. The primary goal of utilizing the LS-SVM is to identify the optimal dataset for training; after that, the dataset is examined to test and validate it using the provided approach. In our prior studies and others, more information about classifiers and LS-SVM applied here are illustrated in

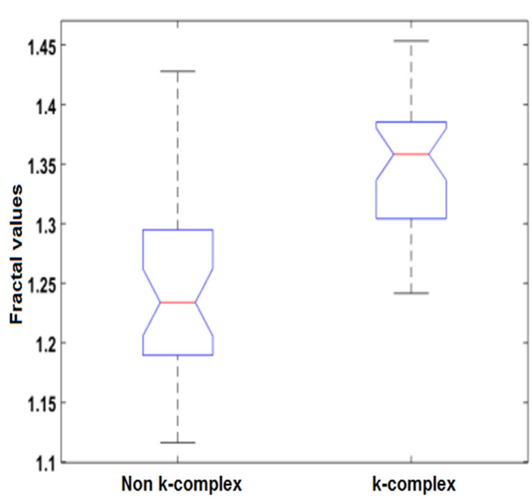
[3,25,29,30,35]. Here, EEG signals *k*-complexes have been characterized using a support vector machine classifier. Classifiers of LS-SVM are the center of several analysts and attention, according to numerous academics. The SL-SMV classifiers, for example, have been proposed by Boser, Guyon, and Vapnik (1992). In order to create a high dimension from the original main data, non-linear mapping was used, and the spearing of the linear optimal hyperplan was investigated after. Several researchers utilized it to categorize EEG signal data, such as the sleep stage and epileptic seizures. The LS-SVM machine has been utilized in this work as the better classifier to separate between the two non and *k*-complexes segments and to contrast their results with other classifiers. In this work, Matlab 2020 is utilized to study the LS-SVM. In order to verify the proposed method, the obtained results were compared with several classifiers (naïve Bayes, artificial neural networks).

5. Experimental results

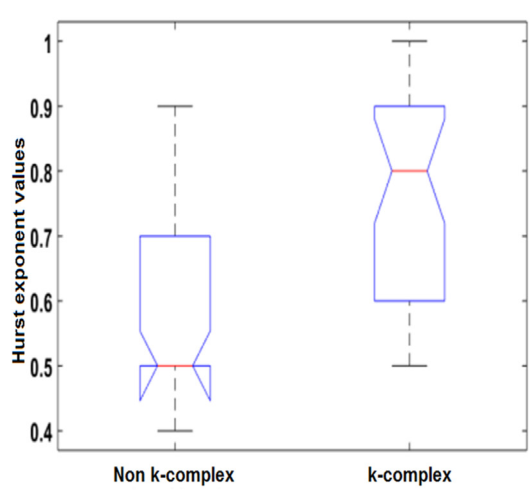
In this study, a number of experiments were used to assess the effectiveness and functioning of our strategy. In Section 2, the datasets used in this study were described. The *k*-complexes have been recognized in the EEG data using a variety of characteristics. Using the sliding window approach, EEG signals were divided into small parts. The window was changed to 0.5 s with 0.4 s of overlap. The entire EEG 0.5 s signal was broken down into 27 main components. These characteristics were taken away while applying the Katz algorithm and the Tunable *Q*-factor wavelet transform. They were first scrutinized before being used as input for the square of the vector machine's classifier support to characterize the discovered biosignal waveforms during the second stage of sleep. The primary intent of choosing a different set of characteristics was to assess how well the characteristics could describe *k*-complexes. More information on the examination of the features will be provided in the following section.

The studies' findings were put to use in Matlab 2020, which required a machine with an Intel ® core i7 (TVM), 3.40 GH processor, and 8.00 GB of RAM. According to the research, selecting the optimum feature set is crucial for categorizing EEG problems. The results showed that our suggested method's classification results are improved by precise feature selection. Our research showed that the ability of the excerpted-different feature set to characterize *k*-complexes was tested. We discovered the traits during the investigation that demonstrated their capacity to recognize the crucial sleep stage 2 characteristics. An illustration of a box plot of some features is shown in Fig. 9. The (non) *k*-complexes FDs are shown in Fig. 9 (a). The fractal dimension values ranged between 1.0 and 2.0. We see that the characteristics of FD could be used to describe all of the (non) *k*-complexes.

The non-linear characteristics were also used in this work because the EEG signals are inherently non-stationary and nonlinear, making classification challenging. 22 characteristics, namely variation, maximum, Hurst exponent, minimum, mean energy, mode, entropy, kurtosis, mean, range, first quartile, and standard deviation, skewness, median, second quartile, time centroid. Shape factor, margin



a) A fractal dimension feature



b) A Hurst exponent (H) feature

Fig. 9. A fractal dimension feature and a Hurst exponent (H) features for identifying k-complexes in EEG signals.

factor, Zero-crossing rate, crest factor, short energy, and impulse factor. In order to distinguish EEG *k*-complexes, these features were merged with fractal and frequency features that were retrieved using Katz's technique and PSD. Our results demonstrate that, when compared to other sub-bands, sub-band 3 (SB3) produced great outcomes during the training phase. The Hurst exponent, for instance, was a significant non-linear property taken from the TQWT. It performed exceptionally well at classifying every *k*-complex present in the EEG recordings. An instance of (non) *k*-complexes plot box of the Hurst exponent is shown in Fig. 9 (b).

The Hurst exponent (H) has values between 0 and 1, where H0.5 denotes a negative correlation between persistent natural series or increments and H > 1 denotes a positive correlation between anti-persistent time series. But as shown in Fig. 9 (b), among all non-complexes and *k*-complexes, the Hurst exponent features provided values with the strongest correlation between 0.5 and 1, and this feature was able to recognize the frequent biosignal waveform in sleep stages 2. The capacity of the nonlinear features to identify *k*-complexes in the EEG data was proved by the H feature, which is a measurement of long-range dependency within a signal. In this study, mean energy and another example, entropy characteristic, were also used. The fact that mean energy was employed to measure a range of activities suggests that different phases of sleep should be investigated. Entropy characteristics are also among the most crucial non-linear features that enable the recognition of regularities in complex waveforms, such as *k*-complexes. The results imply that the *k*-complexes in the EEG data may be described by mean energy and entropy characteristics, as shown in Fig. 10. Skewness and kurtosis were employed as features in this study. Because some EEG waveforms, such as *k*-complexes, have skewed distributions while others have symmetrical distributions, it is imperative to exploit such features. For skewed distributions, these properties were presented as relevant metrics. Our results demonstrate that, as compared to segments that do not contain *k*-complexes, the kurtosis characteristics represent a potential to distinguish the *k*-complexes, while the skewness feature exhibited a negative reflection to do so. As a result, while the skewness feature is not used in this work, it can be employed to distinguish EGG *k*-complexes. Fig. 11 illustrates kurtosis and skewness box plots of (non) *k*-complexes.

The selection of feature technique can choose prominent features while reducing the number of features extracted. This phase is important because it can reduce the number of features our suggested technique must calculate before it can be sent to classifiers. In order to evaluate the features that were extracted and lower the database dimensionality by separating out the crucial features from the unimportant ones, testing computational hypotheses can be a statistically viable strategy.

One of the effective feature selection strategies used in this study to find the relevant features is the one-way analysis of variance of computational hypothesis testing, which is shown in Table 1. Testing

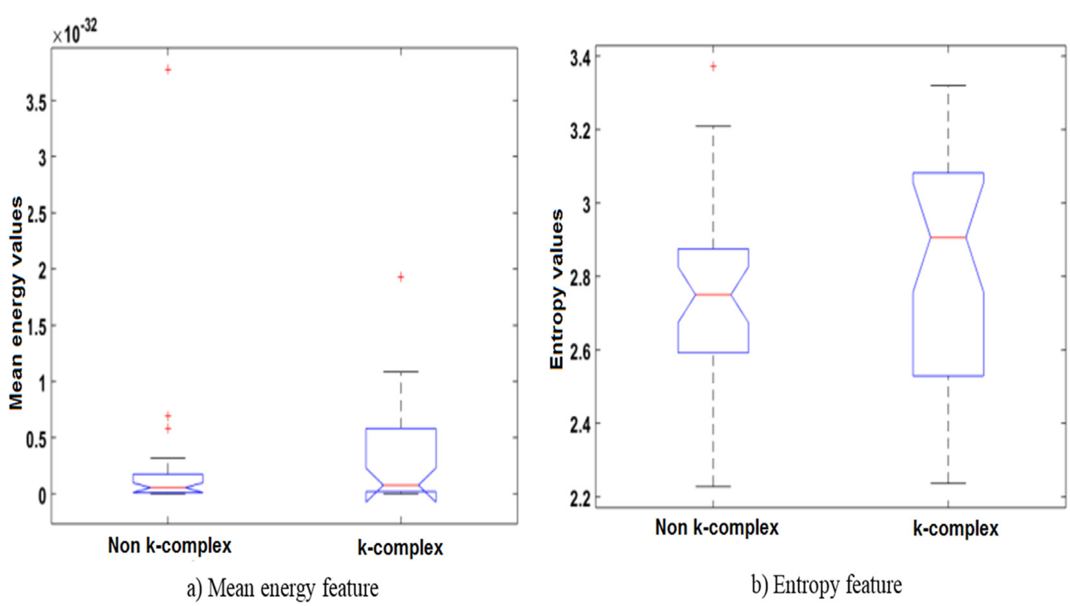
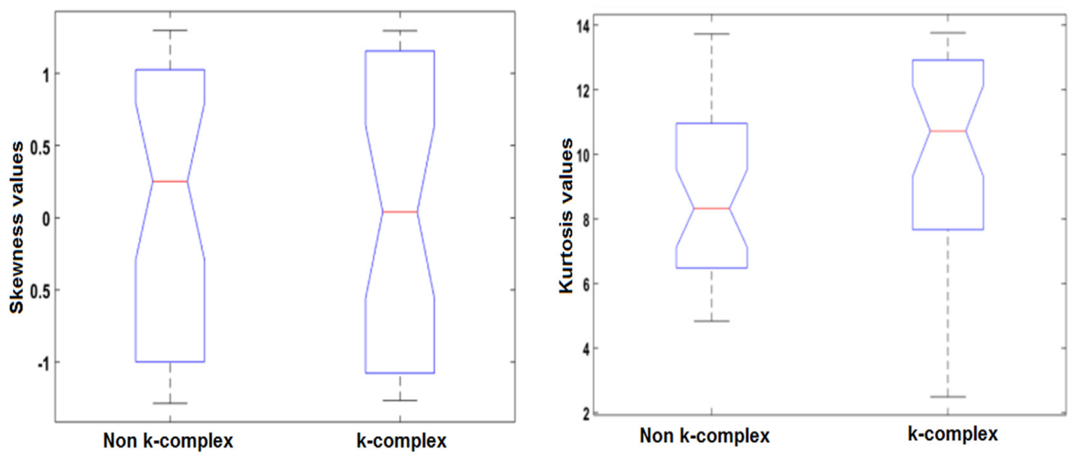


Fig. 10. Mean energy and entropy features to discover k-complexes and non k-complexes in signals of EEG.



a) Skewness features b) Kurtosis feature

Fig. 11. Skewness and Kurtosis features to identify k-complexes and non k-complexes in signals of EEG.

of Computational hypothesis is applied to assess 27 feature sets, variation, maximum, Hurst exponent, minimum, mean energy, mode, entropy, kurtosis, mean, range, first quartile, and standard deviation, skewness, median, second quartile, time centroid. Shape factor, margin factor, zero-crossing rate, crest factor, short energy, and impulse factor were therefore crucial and were chosen to display EEG signals because they received a less than ($p < = 0.05$) P-value. However, because they are not necessary for characterizing the ($P <= 0.05$) *k*-complexes, the characteristics of "crest factor, first quartile, second quartile, skewness, mean, median, median frequency, mode, and mean power" are not chosen. They were therefore used to describe *k*-complexes. Finally, this research used 18 features to characterize *k*-complexes.

Our suggested method's experimental findings are shown. Using a cross-validation approach with 6-fold in the test set, all subject findings were recorded on their specificity accuracy, and sensitivity. In this study, the schema was also evaluated using the Kappa coefficient. For categorical portions, the agreement of the inter-rater metric is more important than the agreement of percent. This is so that Kappa can take into account coincident agreements. The eighteen features were applied to each of the tests in this part in order to determine their functionality. The classification of the suggested algorithm's findings is illustrated in Table 3. The average SVM accuracy, specificity, sensitivity, and kappa coefficient were, respectively, 98.3%, 96.4%, 91.6%, 0.92%, and 0.93%, according to the experiment's results in Table 3. These findings have been compared to the experts' (expert1) scoring in order to determine if our suggested approach and the experts' scoring are both accurate.

5.1. Comparing our proposed with other classifiers

We evaluate our method's performance by contrasting it with other approaches. Two comparisons will be shown here. Our method performance was first contrasted to that of naïve Bayes, artificial neural networks, and *k*-means, these three widely used classification methods. The findings of our method were compared to others using a similar database in the second contrast. The identical features vector was simultaneously used as an input to naive Bayes, LS-SVM, ANN classifiers, and *k*-means for the purpose of comparison with other classifiers. Table 4 provides an illustration of the obtained

Table 3. Proposed method findings of k-complexes detection.

Fold No.	Accuracy%	Sensitivity%	Specificity%	*J*- statistic%	Kappa Coefficient%
Fold 1	98.8	96.8	91.6	0.92	0.92
Fold 2	97.9	96	92.9	0.93	0.93
Fold 3	99	97	93.9	0.95	0.97
Fold 4	98.5	95.8	90.9	0.91	0.91
Fold 5	97	95.8	88.3	0.89	0.93
Fold 6	97	95.8	88.3	0.89	0.93
Average	**98.3**	**96.5**	**91.6**	**0.92**	**0.93**

Table 4. Proposed method results to find k-complexes (average of classification results based on mean ± standard) with using 6-fold cross-validation.

Classifier Types	Accuracy	Sensitivity	Specificity
Naive bayes	82.1 ± 1.70	85.6 ± 1.23	81.8 ± 1.50
ANN	85.5 ± 1.46	82 ± 1.72	79.4 ± 1.64
k-means	92.7 ± 1.92	91 ± 1.66	83.2 ± 1.75
LS-SVM	**98.3 ± 0.47**	**96.5 ± 0.64**	**91.6 ± 0.75**

results. The effectiveness of our approach outperformed that of ANN, *k*-means, and naive Bayes. The specificity, sensitivity, and accuracy of the 6-Fold LS-SVM classifier under cross-validation were 98.3 ± 0.47, 96.5 ± 0.64, and 91.6 ± 0.75, respectively. The *k*-means classifier was used to account for the second-highest result, and ANN produced the lowest result.

5.2. Method performance evaluation by using characteristics of receiver operating (ROCs)

The ROC curve was used to measure the effectiveness of our strategy by applying it to several sets of characteristics. Fig. 12 (a and b) were utilized to demonstrate the ROC includes curve area to our technique utilizing the EEG rather than measuring the computation time required to mimic the classifier of the SVM. Fig. 12 (a) shows our method without feature selection, whereas Fig. 12 (b) shows the effectiveness of our method with feature selection. The greatest observed value of the area under the curve was 0.96 when features were selected, see Fig. 12 (b), and it was roughly 0.78 when features were not selected, see Fig. 12 (a). The outcomes demonstrated the method's growing capability for identifying (non) *k*-complex segments. When the features were analyzed before being assigned to the classifier, it produced excellent results.

5.3. Evaluation of the performance of the proposed model using the Taylor diagram

The Taylor diagram was used to compute the inverse cosine angle to the correlation coefficient, and the performance of the proposed technique employing the LS-SVM classifier was compared to that of other classifiers. Fig. 13 illustrates the comparison of three classifiers utilizing the Taylor diagram. In this work, we employed correlation coefficient and standard deviation to improve the closest fit classifier's capacity to identify *k*-complexes in EEG data.

Fig. 13 shows the results of LS-SVM classifier outperformed other classifiers in terms of generating the highest correlation coefficient value and producing a forecast that was the most accurate with respect to the original data. The gap between the observed and dataset predicted in the datasets testing was therefore cut in half by employing the classifier of LS-SVM with the features of multi-domain. Standard deviations and the greatest correlation coefficient were both 0.90 percent. The extreme learning machine classifier, however, produced the 2nd best outcomes, see Fig. 13. The base results, see Fig. 13, were also attained using an ANN classifier. The classifier of LS-SVM presented an overall improved performance when employed with our technique to differentiate *k*-complexes, including stronger correlation coefficients and a smaller deviation (Fig. 13). Finally, this strategy may offer a practical and accurate way to identify *k*-complexes, which will aid sleep specialists in correctly diagnosing and treating sleep disorders.

5.4. Comparison of our proposed performance employing various features

The efficiency of the recommended technique is assessed by a number of experiments, and the

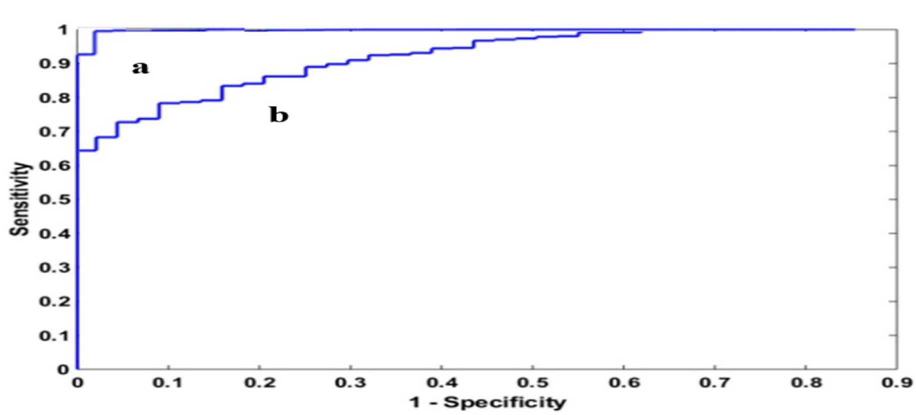


Fig. 12. a and b performance of our method utilizing ROC curve (b) and without selection of features (a).

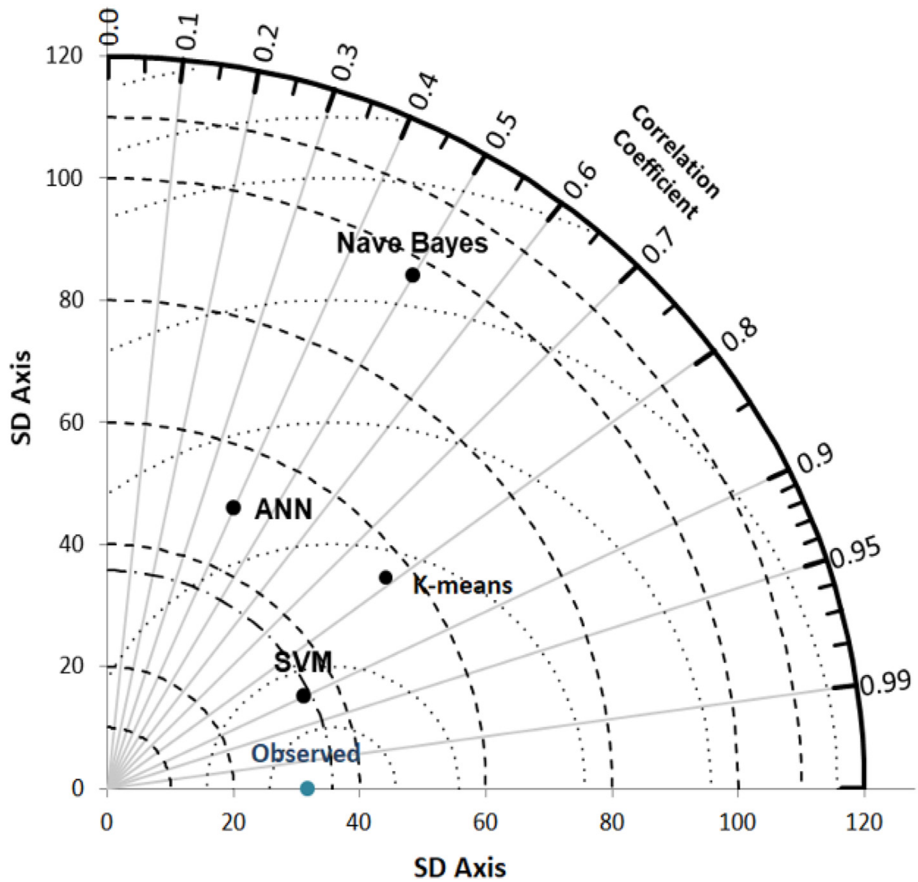


Fig. 13. Taylor diagram on the correlation coefficient used to contrast our method performance and other classifiers.

results are contrasted utilizing a variety of feature types. A specific set of features was applied and utilized with the same methods as in Section 3. The behaviors of these features were examined in order to distinguish the *k*-complexes using a variety of criteria, including fractal, statistics, nonlinear, frequency, and feature components, see Table 5. It was seen that every *k*-complex is grouped according to a particular characteristics set. In order to determine which features are best able to accurately recognize *k*-complexes, the characteristics were taken out from each 0.5 s EEG data and then sent singularly to the classifier of LS-SVM. The classification outcomes of the suggested strategy employing various feature types are displayed in Table 5. This table illustrates the accuracy of the LS-SVM classifiers for the statistic, frequency, fractal, nonlinear, and combination of features was 52.3%, 56.2%, 52.4%, 56%, and 98.3%, respectively. These findings showed that a combination of feature sets boosted the suggested method's classification accuracy rate by 43%.

Table 5. The proposed approach outcomes utilizing various features.

Feature	Range of validation metrics (6-fold)	Average
Statistic features	52%−54%	52.3%
Frequency features	55.1%−57%	56.2%
Fractal features	51%−53.4%	52.4%
Nonlinear features	53%−58.1%	56%
Combination of features	97%−98.8%	98.3%

The effectiveness of the suggested method employing a mixture of features was also assessed in comparison with several sets of features, including statistics, fractals, frequencies, nonlinear features, and a combination of features using the ROC curve. A useful statistic for examining the relationship between sensitivity and specificity is the ROC curve. The outcomes are displayed in Fig. 14. The outcomes in Fig. 14 (a)_(e) demonstrate how precisely *k*-complexes can be represented in EEG by features. Fig. 14 confirms what we discovered in Section 7.1. For the combination of characteristics (multi-domain features), the highest area undervalue of ROC (AUC) was 0.98, whereas the lowest value was 0.51 for features. As a result, it was seen feature combination performance is superior to that of the individual characteristics. As a

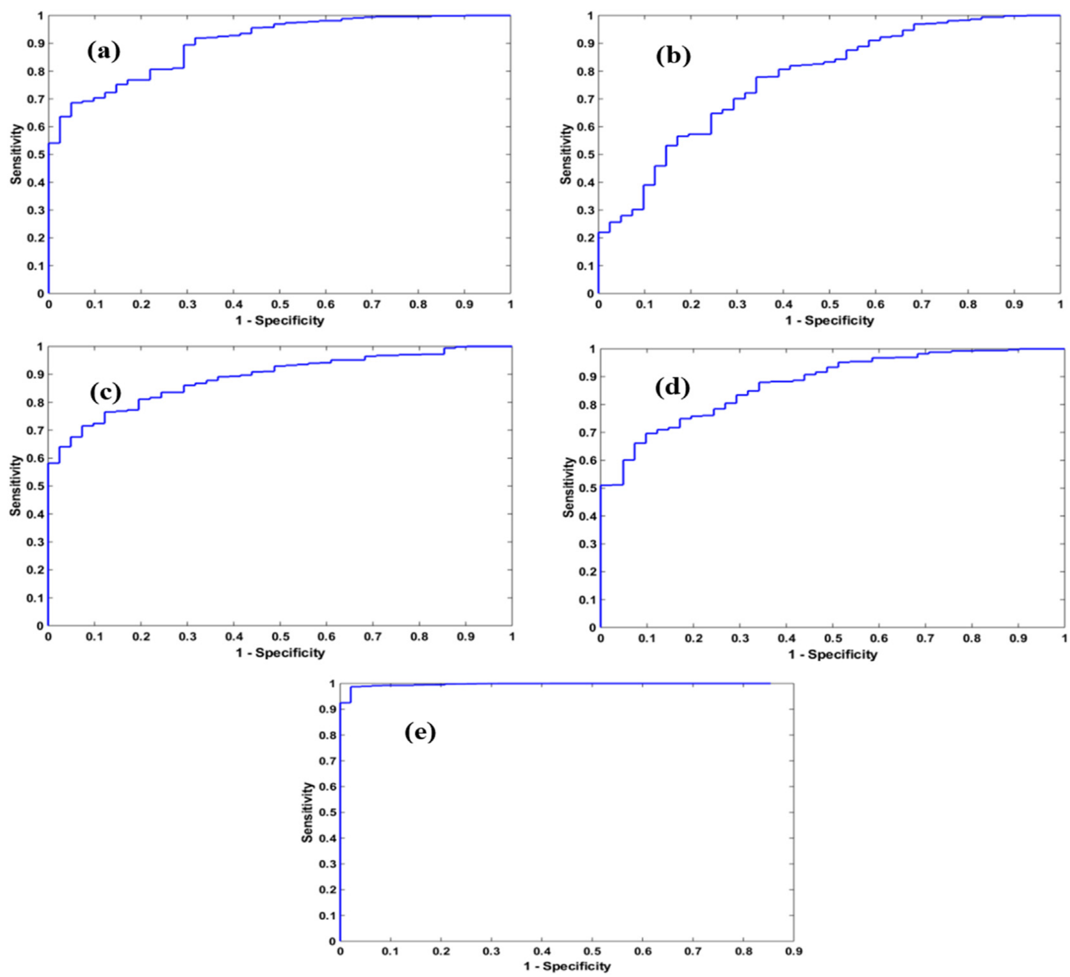


Fig. 14. The ROC curve based on features using LS-SVM classifier: (a) statistical; (b) fractal; (c) frequency, (d) nonlinear and (e) a combination of features.

result, we consider feature combinations in the remaining sections of the study. Finally, the results of the experiments show that our method can identify EEG *k*-complexes data using a combination of features.

5.5. Our suggested method performance when compared to others

Because the EEG data used to determine sleep features in earlier studies varied, it was difficult to evaluate different automatic *k*-complex recognition algorithms. The gained results utilizing the same data must be contrasted in order to be meaningfully compared. In Table 6, the findings from the method categorization were highlighted in bold type. Table 6 provides performance data for a few existing approaches that make use of the same database that we used in our investigation. The outcomes shown in Table 6 show that our method's sensitivity is higher than that of all recently produced approaches. For purposes of comparison, several findings from earlier studies that did not use the same database are also included.

Devuyst et al. [9] suggested a novel method for characterizing *k*-complexes in the EEG data by

Table 6. Existing methods performance comparison with our proposed method.

Used Method	Detection Result
Features extraction using fuzzy thresholds [9]	60.94%
Electrical system of threshold and filtering method [8]	68.0%
Wavelet transformation method [34]	74.0%
Tunable-Q wavelet transform [31]	84.67%
Teager energy operator and WT technique [15]	85.3%
Hjorth parameters and fuzzy decision [37]	86.0%
Artificial neural network based on 14 features [16]	90.0%
k-complex detection's automated methods and signal template [38]	91.2%
Multi-domain feature extraction and selection coupled [4]	92.41%
Using k-complexes with R−CNN and deep transfer learning [39]	92.75%
Proposed method	**98.3%**

applying probability thresholds. In their investigation, they employed a window size and a fuzzy threshold to detect *k*-complexes. This strategy only produced a rate of sensitivity of 60.94. The suggested approach, in contrast, obtained a 0.5 window size and a sensitivity greater than 96.3%. 74% sensitivity was reported by Krohne et al. [36] for the identification of *k*-complexes. They used four features to present their wavelet transformation results. When compared to Krohne et al. [36], our technique produced excellent results. Fuzzy decision-making and Hjorth's parameters were utilized in Migotina et al. [37] to characterize *k*-complexes. The study's reported average sensitivity score of 86% is 11% lower than the approach we suggested in this paper.

Bankman et al. [16] focused on using an artificial neural network classifier to categorize the sleep stage 2 aspects in the EEG signals using 14 features. 90% was the average sensitivity result reported. The sensitivity analysis results fell short of those of our method as well. The *k*-complexes can be described using a method proposed by Erdamar et al. [15]. The researchers were able to attain 85.3% sensitivity for the wavelet transform operator when it was coupled with teager energy. Our approach fared better than the ones suggested by Erdamar et al. [15]. G. Bremer et al. [8] published the findings of their feature extraction method for *k*-complex detection. The electrical system, which used threshold and filtering techniques, was used to extract characteristics from the EEG data before providing the characteristics to the classifier to characterize *k*-complexes. The average sensitivity rate was 68%. In contrast to Bremer et al. [8]'s findings, our strategy produced better outcomes. Table 6 shows that our method is a dependable and practical way to find *k*-complexes.

The suggested approach was further assessed by contrasting it with other research that made use of other datasets. Dorokhov et al. [40] proposed a new method to compare *k*-complexes based on Time-frequency analysis using continuous wavelet transform. In this study, several brain waves for each of the classical frequency ranges were shown using the wavelet spectral power. The P-value was used to assess each retrieved feature. They reported that there are big differences between *k*-complexes of II type in low-frequency bands. Another study was presented by Li and Dong [4], in which multi-domain feature extraction and selection coupled with the RUSBoosted tree model were utilized to detect *k*-complexes. In their work, tunable Q-factor wavelet transform was employed to decompose each EEG segment. An average classification accuracy of 92.41% was recorded. Khasawneh et al. [39] suggested a method to identify *k*-complexes utilizing faster R−CNN and deep transfer learning. The rate of average accuracy was 92.75%. According to the results in Table 6, our proposed method recorded higher results than those existing approaches. Our approach may be utilized as a solid and trustworthy method for the finding of *k*-complexes based on Table 6.

6. Limitations of proposed solution

The suggested method's ability to extract hybrid features effective features from the EEG signal, which is then utilized to recognize waveforms from EEG data, is one of its primary benefits. This method may be applied as an automated platform to assist professionals in identifying anomalies in EEG recordings and aid in their clinical decisions. Furthermore, the suggested technique presents an additional benefit in that it tackles a highly challenging problem: the categorization of EEG signals from the Dream dataset. This difficulty may be resolved by the suggested approach, which for 6-fold cross-validation of the classification yields a classification accuracy rate of 98.3 ± 0.47%. Furthermore, a vast quantity of EEG data may be reduced to a small collection of features that serve as the most representative data points using the suggested feature extraction approach. Furthermore, an additional benefit of the suggested approach, which is predicated on the categorization of EEG signals in waveforms from the Dream dataset, is an extremely challenging task. The suggested technique resolves this issue by achieving a 6-fold cross-validation classification accuracy rate of 98.3 ± 0.47%. Furthermore, a large quantity of EEG data may be reduced to a small collection of features that serve as the best representative data points using the suggested feature extraction approach. Because of its efficient data reduction, this study can handle large EEG datasets at a lower processing cost than previous approaches.

However, the suggested approach has several drawbacks. As can be shown from the experimental findings in Tables 4 and 5, the approach performs very well when applied to an LS-SVM classifier, mediumly well when used to a Naïve Bayes classifier, and poorly when applied to an ANN classifier due to its inflexibility. The fact that the suggested solution requires a lot of computing power and that we only employed a tiny database are other drawbacks. Furthermore, the suggested method's performance did not yield good outcomes for distinct aspects like statistics and frequency features. In the future, we also want to use real databases to recognize waveforms from EEG signals and evaluate our approach with more than two databases. Two further drawbacks of the suggested approach are its high processing cost and the tiny database we employed. Furthermore, specific characteristics like statistics, non-linearity, and frequency features did not yield good performance outcomes from the suggested strategy. Furthermore, we want to employ real databases in the future to evaluate our method's ability to recognize waveforms from EEG signals, as well as several databases.

7. Conclusion

We pointed out here a potentially innovative approach to identify *k*-complexes in EEG data. This approach uses an LS-SVM classifier together with a range of feature sets. Each EEG segment lasted 0.5 s and included an extract of 27 sets of features. We then examined those characteristics before providing them to the classifier. Computational hypothesis testing was used to identify the resilient qualities and strong features in order to identify the EEG *k*-complexes. The EEG signals are inherently non-stationary and nonlinear, making classification challenging, non-linear characteristics, fractal dimension, and frequency features were used in this work to achieve a high level of accuracy rate as well as to make the detection phase of waveforms easy. 18 features out of 27 were selected based on the ANOVA F-test, as shown in Fig. 5 and Table 1. They were employed as the classifier input for LS-SVM in order to describe EEG *k*-complexes. Our experiment's results demonstrated that the collection of 18 characteristics yielded better results, with average specificity, sensitivity, and accuracy values of 91.6%, 98.3%, and 96.5%. In order to check our method's efficiency, we compared it to other techniques and classifiers to assess its potential for describing *k*-complexes. We saw that our technique produced high classification results when compared to other classifier methods. Therefore, the proposed schema could be able to help healthcare providers correctly identify and treat sleep problems. Our method works with a wide range of medical data. It can be applied in the future to emphasize different characteristics of the sleep state, such as slow and vertex waves.

Ethical statement

Not Applicable.

Funding

This work did not receive any external funding.

Acknowledgment

We would like to thank everyone who contributed and worked hard to achieve this work.

References

[1] J.P. Kiley, M.J. Twery, G.H. Gibbons, The National center on sleep disorders Research—Progress and promise, Sleep 42 (6) (2019), https://doi.org/10.1093/sleep/zsz105.

[2] T. Ngo, P.N. Pathirana, M.K. Horne, L.A. Corben, I.H. Harding, D.J. Szmulewicz, Technological evolution in the instrumentation of ataxia severity measurement, IEEE Access 11 (2023) 14006−14027, https://doi.org/10.1109/ACCESS.2023.3243178.

[3] T.F. Zaidi, O. Farooq, EEG sub-bands based sleep stages classification using Fourier Synchrosqueezed transform features, Expert Syst. Appl. 212 (2023) 118752, https://doi.org/10.1016/j.eswa.2022.118752.

[4] Y. Li, X. Dong, A rusboosted tree method for *k*-complex detection using tunable *Q*-factor wavelet transform and multi-domain feature extraction, Front Neurosci. 17 (2023) 1108059, https://doi.org/10.3389/fnins.2023.1108059.

[5] Q. Li, K.K. Ng, C. Simon, C.Y. Yiu, M. Lyu, Recognising situation awareness associated with different workloads using EEG and eye-tracking features in air traffic control tasks, Knowl Base Syst 260 (2023) 110179, https://doi.org/10.1016/j.knosys.2022.110179.

[6] G. Jorgensen, C. Downey, J. Goldin, K. Melehan, P. Rochford, W. Ruehland, An Australasian commentary on the AASM manual for the scoring of sleep and associated Events, Sleep Biol Rhythm 18 (2020) 163−185, https://doi.org/10.1007/s41105-020-00259-9.

[7] T.A. Camilleri, K.P. Camilleri, S.G. Fabri, Automatic detection of spindles and *k*-complexes in sleep EEG using switching multiple models, Biomed Signal Process Control 10 (2014) 117−127, https://doi.org/10.1016/j.bspc.2014.01.010.

[8] G. Bremer, J.R. Smith, I. Karacan, Automatic detection of the *k*-complex in sleep electroencephalograms, in: IEEE Transactions on Biomedical Engineering, 4, 1970, pp. 314−323, https://doi.org/10.1109/TBME.1970.4502759.

[9] S. Devuyst, T. Dutoit, P. Stenuit, M. Kerkhofs, Automatic *k*-complexes detection in sleep EEG recordings using likelihood thresholds, in: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, IEEE, 2010, pp. 4658−4661, https://doi.org/10.1109/IEMBS.2010.5626447.

[10] A.R. Hassan, S. Siuly, Y. Zhang, Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating, Comput Methods Progr Biomed 137 (2016) 247−259, https://doi.org/10.1016/j.cmpb.2016.09.008.

[11] C. Altõn, O. Er, Comparison of different time and frequency domain feature extraction methods on elbow gesture's EMG, European J Interdisc Studies. 2 (2016) 35−44, https://doi.org/10.26417/ejis.v2i3.p35-44.

[12] A. Da Rosa, B. Kemp, T. Paiva, F.L. da Silva, H. Kamphuisen, A model-based detector of vertex waves and *k* complexes in sleep electroencephalogram, Electroencephalogr Clin Neurophysiol 78 (1991) 71−79, https://doi.org/10.1016/0013-4694(91)90021-U.

[13] B.H. Jansen, P.R. Desai, *k*-complex detection using multilayer perceptrons and recurrent networks, Int J Bio Med Comput 37 (1994) 249−257, https://doi.org/10.1016/0020-7101(94)90123-6.

[14] A. Kam, A. Cohen, A. Geva, A. Tarasiuk, Detection of *k*-complexes in sleep EEG using CD-HMM, in: The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 1, 2004, pp. 33−36, https://doi.org/10.1109/IEMBS.2004.1403083.

[15] A. Erdamar, F. Duman, S. Yetkin, A wavelet and teager energy operator based method for automatic detection of *k*-

complex in sleep EEG, Expert Syst Appl 39 (2012) 1284−1290, https://doi.org/10.1016/j.eswa.2011.07.138.
[16] I.N. Bankman, V.G. Sigillito, R.A. Wise, P.L. Smith, Feature-based detection of the *k*-complex wave in the human electroencephalogram using neural networks, IEEE Trans Biomed Eng 39 (1992) 1305−1310, https://doi.org/10.1109/10.184707.
[17] E. Hernández-Pereira, V. Bolón-Canedo, N. Sánchez-Maroño, D. Álvarez-Estévez, V. Moret-Bonillo, A. Alonso-Betanzos, A comparison of performance of *k*-complex classification methods using feature selection, Inf Sci 328 (2016) 1−14, https://doi.org/10.1016/j.ins.2015.08.022.
[18] H. Al-Hadeethi, I. Al-Hadeethi, Using mathematical techniques to analyse biomedical data: A *k*-complexes EEG signal classification study, J Port Sci Res 7 (2024) 208−218, https://doi.org/10.36371/port.2024.special.19.
[19] T. Lajnef, S. Chaibi, J.-B. Eichenlaub, P.M. Ruby, P.-E. Aguera, M. Samet, A. Kachouri, K. Jerbi, Sleep spindle and *k*-complex detection using tunable *Q*-factor wavelet transform and morphological component analysis, Front Human Neuro-sci 9 (2015) 414, https://doi.org/10.3389/fnhum.2015.00414.
[20] A. Parekh, I.W. Selesnick, D.M. Rapoport, I. Ayappa, Detection of *k*-complexes and sleep spindles (DETOKS) using sparse optimization, J Neurosci Methods 251 (2015) 37−46, https://doi.org/10.1016/j.jneumeth.2015.04.006.
[21] C. Dumitrescu, Automatic detection of *k*-complexes using deep neural networks with EEG signals, Prime Archives in Sensors (2022), https://doi.org/10.3390/s21217230.
[22] S. Chambon, V. Thorey, P.J. Arnal, E. Mignot, A. Gramfort, A deep learning architecture to detect events in EEG signals during sleep, in: 2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP), IEEE, 2018, pp. 1−6, https://doi.org/10.1109/MLSP.2018.8517067.
[23] G.H. Oliveira, L.R. Coutinho, J.C. da Silva, I.J. Pinto, J.M. Ferreira, F.J. Silva, D.V. Santos, A.S. Teles, Multitaper-based method for automatic *k*-complex detection in human sleep EEG, Expert Syst Appl 151 (2020) 113331, https://doi.org/10.1016/j.eswa.2020.113331.
[24] M. Gala, J. Mohylova, Detection of *k*-complex in the EEG signal, in: World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany: Vol. 25/4 Image Processing, Biosignal Processing, Modelling and Simulation, Biomechanics, Springer. 2010, pp. 1170−1173, https://doi.org/10.1007/978-3-642-03882-2_311.
[25] Y. Li, P.P. Wen, Clustering technique-based least square support vector machine for EEG signal classification, Comput Methods Progr Biomed 104 (2011) 358−372, https://doi.org/10.1016/j.cmpb.2010.11.014.
[26] S. Ashokkumar, M. Premkumar, S. Anupallavi, V. Jeevanantham, G. Mohanbabu, A. Selvapandian, Quantifying functional connectivity network through synchronization and graph theory approaches for seizure prediction, Wireless Pers Commun (2023) 1−34, https://doi.org/10.1007/s11277-022-10154-w.
[27] M.H. Kabir, S. Mahmood, A. Al Shiam, A.S. Musa Miah, J. Shin, M.K.I. Molla, Investigating feature selection techniques to enhance the performance of EEG-based motor imagery tasks classification, Mathematics 11 (2023) 1921, https://doi.org/10.3390/math11081921.
[28] M.B. Musthafa, S. Huda, Y. Kodera, M.A. Ali, S. Araki, J. Mwaura, Y. Nogami, Optimizing IoT intrusion detection using balanced class distribution, feature selection, and ensemble machine learning techniques, Sensors 24 (2024) 4293, https://doi.org/10.3390/s24134293.
[29] X. Gao, G. Li, A KNN model based on manhattan distance to identify the SNARE proteins, IEEE Access 8 (2020) 112922−112931, https://doi.org/10.1109/ACCESS.2020.3003086.
[30] T.K.K. Ho, N. Armanfard, Self-supervised learning for anomalous channel detection in EEG graphs: Application to seizure analysis, Proc AAAI Conf Artif Intell 37 (2023) 7866−7874, https://doi.org/10.1609/aaai.v37i7.25952.
[31] A. Bhattacharyya, R.B. Pachori, U.R. Acharya, Tunable-Q wavelet transform based multivariate sub-band fuzzy entropy with application to focal EEG signal analysis, Entropy 19 (2017) 99, https://doi.org/10.3390/e19030099.
[32] H.R. Al Ghayab, Y. Li, S. Siuly, S. Abdulla, A feature extraction technique based on tunable *Q*-factor wavelet transform for brain signal classification, J Neurosci Methods 312 (2019) 43−52, https://doi.org/10.1016/j.jneumeth.2018.11.014.
[33] C. Wei, H. Wang, B. Zhou, N. Feng, F. Hu, Y. Lu, D. Jiang, Z. Wang, sEMG signal-based lower limb movements recognition using tunable Q-factor wavelet transform and kraskov entropy, IRBM 44 (2023) 100773, https://doi.org/10.1016/j.irbm.2023.100773.
[34] A. Kamble, P.H. Ghare, V. Kumar, Optimized rational dilation wavelet transform for automatic imagined speech recognition, in: IEEE Transactions on Instrumentation and Measurement, 72, 2023, pp. 1−10, https://doi.org/10.1109/TIM.2023.3241973.
[35] E. Tuncer, E.D. Bolat, Channel based epilepsy seizure type detection from electroencephalography (EEG) signals with machine learning techniques, Biocybern Biomed Eng 42 (2022) 575−595, https://doi.org/10.1016/j.bbe.2022.04.004.
[36] L.K. Krohne, R.B. Hansen, J.A. Christensen, H.B. Sorensen, P. Jennum, Detection of *k*-complexes based on the wavelet transform, in: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2014, pp. 5450−5453, https://doi.org/10.1109/EMBC.2014.6944859.
[37] D. Migotina, A. Rosa, A. Fred, Automatic *k*-complex detection using hjorth parameters and fuzzy decision, in: Proceedings of the 2010 ACM Symposium on Applied Computing, 2010, pp. 979−980, https://doi.org/10.1145/1774088.1774293.
[38] R. Ranjan, B.C. Sahana, Automatic identification of *k*-complex in sleep EEG signal for clinical practice, in: 2022 IEEE Delhi Section Conference (DELCON), IEEE. 2022, pp. 1−8, https://doi.org/10.1109/DELCON54057.2022.9753300.
[39] N. Khasawneh, M. Fraiwan, L. Fraiwan, Detection of *k*-complexes in EEG waveform images using faster R-CNN and deep transfer learning, BMC Med Inf Decis Making 22 (2022) 1−14, https://doi.org/10.1186/s12911-022-02042-x.
[40] V. Dorokhov, A. Runnova, O. Tkachenko, A. Taranov, G. Arseniev, A. Kiselev, A. Selskii, A. Orlova, M. Zhuravlev, Analysis two types of *k* complexes on the human EEG based on classical continuous wavelet transform, Chaos: An Interdis J Nonlinear Sci 33 (2023) 3, https://doi.org/10.1063/5.0143284.