Statistical-Hypothesis-Aided Tests for Epilepsy Classification

Alaa Alqatawneh [1,†], Rania Alhalaseh [1,*,†], Ahmad Hassanat [2,†] and Mohammad Abbadi [1,†]

1   Computer Science Department, Mutah University, Karak 61710, Jordan; alaaqatawneh@hotmail.com (A.A.); abbadi@mutah.edu.jo (M.A.)
2   Computer Department, Community College University of Tabuk, Tabuk 71491, Saudi Arabia; hasanat@mutah.edu.jo
*   Correspondence: halaseh@mutah.edu.jo
†   These authors contributed equally to this work.

check for updates

Abstract: In this paper, an efficient, accurate, and nonparametric epilepsy detection and classification approach based on electroencephalogram (EEG) signals is proposed. The proposed approach mainly depends on a feature extraction process that is conducted using a set of statistical tests. Among the many existing tests, those fit with processed data and for the purpose of the proposed approach were used. From each test, various output scalars were extracted and used as features in the proposed detection and classification task. Experiments that were conducted on the basis of a Bonn University dataset showed that the proposed approach had very accurate results (98.4%) in the detection task and outperformed state-of-the-art methods in a similar task on the same dataset. The proposed approach also had accurate results (94.0%) in the classification task, but it did not outperform state-of-the-art methods in a similar task on the same dataset. However, the proposed approach had less time complexity in comparison with those methods that achieved better results.

Keywords: biomedical signal processing; electromyography; multiple-signal processing; EEG; machine learning; epilepsy

1. Introduction

Epilepsy is a brain disorder that affects the whole nervous system and is characterized by high-frequency and high-voltage brain waves called seizures. An epileptic seizure is defined as a transient symptom of excessive or synchronous neuronal activity in the brain [1]. Epileptic seizures have bad effects on the patient, ranging from attention lapses to muscle jerks. According to a World Health Organization [2] report, there are around 50 million epilepsy patients worldwide. Early diagnosis of epilepsy helps identify suitable precautions.

Epilepsy is assessed similarly to many other brain disorders by electroencephalogram (EEG) [3,4]. EEG is a noninvasive, low-cost, well-established, and reliable technique that is used for diagnosing brain-related disorders such as epilepsy, tumors, and depression. EEG captures potential differences in the brain using electrodes that pick up signals and send them to the main EEG machine, which saves the generated signals.

For epilepsy patients, there are two phases that vary in diagnosis and in the signals that are captured by the EEG, namely the ictal and interictal phases. The ictal phase represents the seizure and is characterized by abnormal signals, while the interictal phase is the intermediate phase between seizures that contains different forms of signals. These phases have different characteristics compared to the signals of healthy subjects. Thus, EEG data can be used to detect and classify epilepsy disorders [3]. Figure 1 illustrates the three different states.
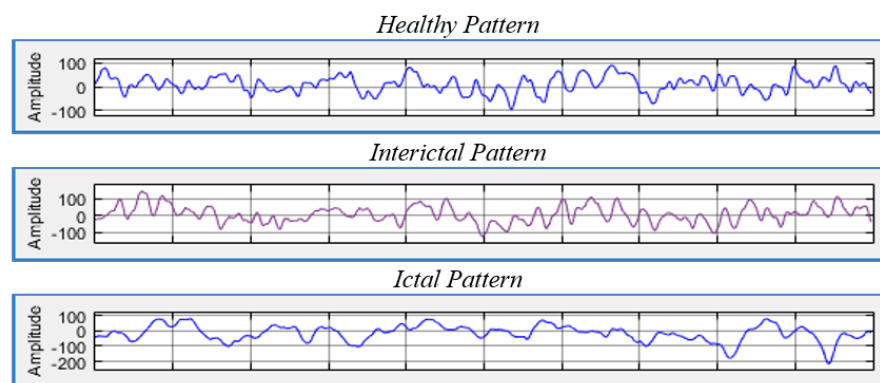
Figure 1. Brain waves present in different diagnoses.

The detection of seizure based on EEG signals can save millions of lives, but the analysis of such nonlinear and nonstationary signals is anything but trivial. EEG signals have different patterns depending on the person and their state (e.g., awake, sleep, and alert). Thus, EEG signal analysis requires consideration of the expected patterns in these signals. To detect seizures, ictal signals need to be identified and differentiated from interictal and healthy subject signals. As the first step in EEG analysis, features are extracted from underlying recorded signals. Features can be any compact representation of the given signal. However, features that are able to characterize epileptic and nonepileptic states should have high discriminative and interassociation abilities. Discriminative ability refers to the ability to give different representation to various classes (e.g., healthy, ictal, and interictal), and different values that do not overlap. Interassociation refers to the ability to give a similar representation for each class with similar or identical values [5]. The feature extraction time is also critical, given that the signal to be processed is huge, and applications of epilepsy detection, especially for seizure detections, occur in real time. Thus, the complexity of the feature extraction process should be as low as possible.

In this paper, an efficient, accurate, and nonparametric epilepsy detection and classification approach based on EEG signals is proposed. Epilepsy detection and classification is based on two components: extracted features and the used classification algorithm. The use of machine learning and its algorithms is quite promising, where different approaches have not yet been investigated. Furthermore, the accuracy of EEG signal classification and detection of seizure stages mainly depends on the used features. In this paper, selected statistical tests were used as the basis for extracting highly discriminating and time-efficient features, which were then used in epilepsy detection and classification.

2. Related Work

Epilepsy seizure detection, identification, and classification are implemented on the basis of feature extraction methods being applied to EEG signals. Various methods with various features were proposed in the literature, and these can be classified into three categories: nonlinear, Fourier-based, and wavelet-based detection methods.

2.1. Nonlinear Feature-Based Epilepsy Seizure Detection

Kannathal et al. [6] used spectral (Shannon), Renyi, Kalmogorov–Sinai, and approximate entropies with an adaptive neurofuzzy classifier (AFC) to only differentiate between two classes, namely the epilepsy ictal state and normal EEG signals. Approximate, spectral, and Kalmogorov–Sinai entropies gave higher values for the normal signals compared with the values obtained from epilepsy patients. For Renyi entropy, it was noted that the values of epilepsy patients and normal subjects were identical. Accordingly, the discriminating abilities of these features were ranked from good to bad. According to the experiments conducted using part of the Bonn University dataset, which includes 200 sample

signals for the normal class and 100 sample signals for the ictal class, the results of these features were 90% accurate.

Acharya et al. [5] used a correlation dimension, a fractal dimension, the Hurst exponent, the largest Lyapunov exponent, and approximate entropy with Support Vector Machine (SVM) and Gaussian Mixture Model (GMM) classifiers. Similarly, Acharya et al. [7], Yang et al. [8], S. Vijith et al. [9], Thilagaraj et al. [10], and Li et al. [11] used similar or other features and classifiers and had similar or slightly more accurate results. Overall, the use of nonlinear features for epilepsy detection and classification resulted in an output with an accuracy of up to 98% using features that required time with a complexity of $O(n \log n)$ or higher, where n is the length of the acquired signal. This complexity prevents epilepsy detection being implemented in real time.

2.2. Fourier-Based Epilepsy Seizure Detection

Discrete Fourier Transformation (DFT) is used to capture the frequency changes in the time domain and construct the so-called frequency space. The problem with DFT is its inability to capture the sensitivity of nonstationary EEG signals. This is because DFT transfers the signal by summing up the frequency components of infinite duration. Thus, DFT cannot be used to capture the events in the time series, and it is not able to distinguish among different EEG patterns. Accordingly, with EEG-based epilepsy detection, a new extension of Fourier called short-time Fourier transformation (STFT) is used in order to overcome the limitations of Fourier transforms in dealing with sensitive EEG signals [12].

Nijsen et al. [12] implemented STFT and extracted STFT coefficients, coefficient range, and normalized coefficients. The sensitivity of these features was statistically analyzed on the basis of real data obtained from seven patients, each with two data samples of at least one seizure state and a duration of five minutes. Experiments showed that the sensitivity of these features ranged from 0.66 to 1. Krishnakumar and Thanushkodi [13] implemented STFT as a preprocessing step with Independent Component Analysis (ICA). Then, three linear and nonlinear features, standard deviation, correlation dimension, and Lyapunov exponents, were extracted and used with an Artificial Neural Network (ANN) classifier, which had an accuracy of 96.2%.

Kovcs et al. [14] and Samiee et al. [15] used similar features with an ANN classifier and the Bonn dataset and achieved an accuracy of 98.1%. Szuflitowska and Orlowski [16] implemented a similar approach and obtained an accuracy of 85%.

Overall, the use of STFT for epilepsy detection and classification either uses measurements of coefficients, such as the maximum, minimum, and average, or STFT, which is used as a preprocessing step followed by nonlinear feature extraction. Similar to other work on epilepsy detection, the results of the reviewed literature showed that STFT gives accurate results with a minimum complexity of $O(n \log n)$.

2.3. Wavelet-Based Epilepsy Seizure Detection

Übeyli et al. [17] extracted the maximum, minimum, mean, and standard deviation of wavelet output functions after applying fourth-order wavelets. The implemented classification task differentiated between five data classes: the epilepsy ictal state and normal with open and closed eyes and the interictal state from opposite hippocampal locations. The results of the conducted experiments using part of the Bonn University dataset had an accuracy of 98%. Sadati et al. [18] used the output functions of implementing five levels of wavelet filters with adaptive neurofuzzy and SVM classifiers to only differentiate between two classes, namely the epilepsy ictal state and normal EEG signals. The results of the conducted experiments using the Bonn University dataset, which included 100 sample signals for the normal and ictal states, had an accuracy of 85.9% for the fuzzy classifier and 83.1% for the SVM.

Similarly, Jahankhani et al. [19], Subasi [20], Subasi [21], Costa et al. [22], Guo et al. [23], and Orhan et al. [24] used wavelets for epilepsy detection and classification. Similar to STFT, the wavelet-based approaches either used the coefficients themselves or related measurements such as

the maximum, minimum, and average. The results of the reviewed literature showed that the wavelets had accurate results with a minimum complexity of $O(n \log n)$.

Another study by Khan et al. [25] proposed a method for the automatic detection of seizure onset, in which detection is based on two statistical features of the signals, skewness and kurtosis, with a wavelet-based feature normalized with a coefficient variation that is extracted from the data. The approach was shown to detect all seizures in the experiment with an average latency period of 3.2 s.

The previously discussed literature shows that there are common and unique issues regarding seizure detection approaches that can be summarized in four points. First, a variety of classifiers were used in the literature. Most of the experiments conducted for epilepsy detection have generally used different classification algorithms. Second, various features were included in the classification process. Third, the dataset provided by Bonn University has commonly been used in experiments because it is the only publicly available dataset for epilepsy. Fourth, the feature extraction step is time-consuming.

Although the accuracy of some approaches reported in the literature was above 98%, not all the implemented approaches based on the the Bonn University dataset as a whole were conducted to differentiate between the three classes of data, namely the ictal, interictal, and normal states. Moreover, the complexity of the state-of-the-art methods is high for real-time processing. Accordingly, there is a need to enhance the accuracy of epilepsy detection and classification and reduce the complexity of the used approaches.

3. Proposed Work

The proposed approach uses a statistical test that refers back to the data distribution of EEG signals. In general, data can be distributed in different ways, such as spread left, spread right, jumbled up, or centralized. Centralized distribution naturally occurs in many situations, including diagnosis signals using EEG. Centralized distribution with a bell shape is called normal distribution. In normal distribution, the probability density is higher in the middle compared to the tails on the right and left of the center. The bell shape of a normal distribution is fixed, but it stretches and shifts on the basis of mean and standard data deviation. Having a normal distribution is important because the Central Limit Theorem (CLT) states that independent random variable sums tend to follow a normal distribution in most cases.

Furthermore, EEG signals for epilepsy diagnosis can be assumed to follow a normal distribution to some extent, given the CLT. However, due to various effects on the data, such as noise, the biological characteristics of the scanned subject, and surrounding effects, it is impossible to have the same distribution even for the same subject in the same state. Nevertheless, as data tend to have specific characteristics in each state, they are assumed to follow specific patterns in terms of diverging from or approaching a normal distribution. This assumption is the motivation of the proposed work. The EEG signals have specific characteristics in each state, and assumed to follow a specific pattern in this paper.

Figure 2 presents a block diagram of the framework proposed in this paper, which can be summarized by the following steps:

- Step 1: Signal preprocessing:
  The signal is initially preprocessed to remove the effects of any artifacts and noise in the data.
- Step 2: Probability Density Function (PDF) fitting:
  A PDF fitting method is performed.
- Step 3: Hypothesis tests:
  The resulting PDF function with the preprocessed data is used as the input to a set of statistical hypothesis tests.
- Step 4: Machine-learning methods:
  The output of the hypothesis tests forms a structured dataset that is employed in machine-learning algorithms, including feature selection and classification methods.

Furthermore, the implemented work is constructed on the basis of the three following main techniques:

1. The PDF fitting method, which is responsible for forming an associated normal distribution, is applied;
2. The divergence between the input data and associated distribution is calculated using statistical (hypothesis) tests; and
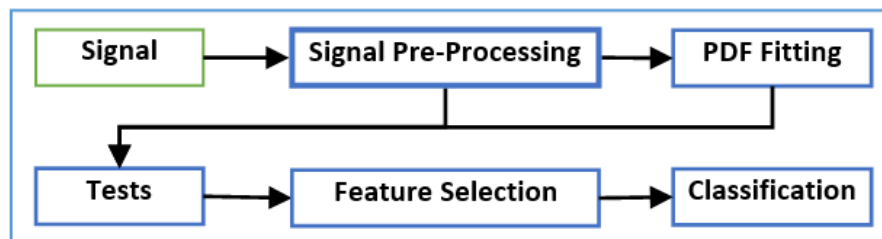3. The machine-learning algorithm is employed for classification and detection purposes.

Figure 2. Proposed framework.

In the following subsections, the steps shown in Figure 2 are discussed in detail.

3.1. Step 1: Signal Preprocessing

EEG signals are affected by various types of interference and artifacts that modify the original signals generated in the brain. The band-pass filter, which was proven to eliminate noise in EEG signals, is used to remove artifacts and noise in the input signal. The band-pass filter allows signals within a frequency range to pass as is while blocking anything outside that range Rioul and Vetterli [26], as illustrated in Figure 3. As the signals are filtered, they are then represented as time-series vectors $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$, recorded at various time instances from 1 to n. Finally, the data are represented in a histogram (values and frequencies).
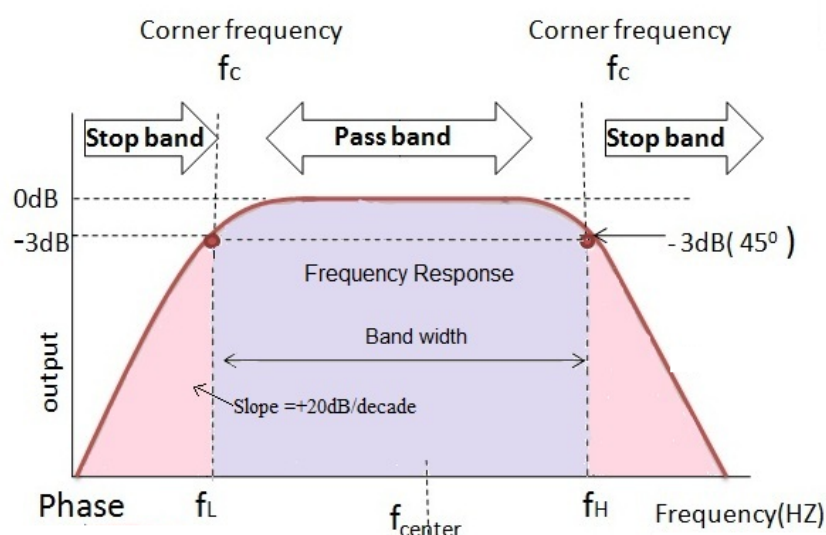

Figure 3. Band-pass filter [26].

3.2. Step 2: PDF Fitting

The PDF is estimated for each sample datum by calculating the probability of each bin in the histogram and then calculating the statistics that summarize the density. The density estimation

process is responsible for constructing the unobserved PDF from the observed data. The first step in density estimation is to determine the type of modeled variable, i.e., continuous vs. discrete, and to select a distribution type that can represent the variable and fit the data characteristics. Because the preprocessed input signals are discrete, they can be represented using binomial and uniform discrete distributions if the data are symmetric (those distributed normally on the side or the central value), while geometric, negative bidirectional, and hygrometric distributions are used if the data are asymmetric. A normal distribution was selected to model the EEG signal. This is because a normal distribution often occurs with any type of data. According to the CLT, for any independent random variable, the normalized sum of the values is distributed according to a normal distribution. This is a key concept in probability theory because it implies that a normal distribution can be applied to any problem including variables generated using other distributions.

In addition to the CLT, there are further reasons that support the selection of a normal distribution in this work.

- A normal distribution is characterized by two parameters, the mean μ and standard deviation σ. Therefore, the fitting process for such distribution is relatively uncomplicated and does not consume a significant amount of processing time.
- A normal distribution is used with data that tend to take a central value. In EEG signals, it is assumed that signal values are centralized at a specific wave type. In other words, a normal distribution is used with data that tend to have equal positive and negative values from the central value, which is the case in EEG signals. Although this assumption might be weak, it is supported by two facts. First, the distribution itself is used to model the data and not the features. Second, the normality assumption is evaluated at each piece of sample data in the next step. Other distributions require harsh assumptions to be made, which cannot be risked in the proposed approach.

The second step of PDF fitting is to generate the distribution function curve using the input data. The normal function is generated by using the value histogram and their frequencies based on the calculation of the mean μ and standard deviation σ. After this process, the data are represented by two vectors, the preprocessed data vector **x**, and the normal distribution mode μ, σ. Finally, in order to check the normality of the data, a Chi-square variance test is used individually with each of the items of sample data to check the normality of the data with $\alpha = 0.05$.

3.3. Step 3: Hypothesis Tests

After the PDF is estimated for each observed data sample, a set of hypothesis tests are implemented. The purpose of each implemented test depends on the type of the test itself; this includes the quality of fitting between the data and the PDF, the data randomness, and the data correlation. Hypothesis tests are a mechanism for making decisions about the independent random variable that generates the data, which can be a process, a natural measure, etc. Any hypothesis test is implemented to statistically determine whether there is enough evidence to reject the null hypothesis or not. Accordingly, the output of any statistical test is the hypothesis variable h, which is either accepted or rejected. Commonly, an accepted h is indicated by 0 and a rejected h is indicated by 1.

The probability value (p-value) is a measurement of the test accuracy and it is used to determine whether or not to reject the null hypothesis H_0. The larger the p-value is , the more random the results are, which indicates the inability to reject H_0. The p-value is usually mapped from the output of the test with the aid of a table that matches between the output of the test calculation and the p-value. Then, in order to determine whether the p-value is considered large or small, a confidence level or critical value α is used to determine the threshold between large and small probabilities. Hence, if the produced p-value is less than α, H_0 can be rejected with confidence $1 - \alpha$; on the other hand, if the p-value is equal to or greater than α, H_0 cannot be rejected.

Accordingly, various tests with a predetermined null hypothesis and alternative hypothesis are used, and they can be categorized into three groups: distribution tests, location tests, and dispersion

tests. Of the three groups, selected tests that fit with the processed data and the purpose of the proposed approach are implemented. Some statistical tests, such as the *t*-test, are not useful, because they are used to test whether the population mean is equal to a specified value; in our case, the mean is calculated from the data themselves, so it is always true. For each test, the focus of the proposed work is neither to accept nor to reject the null hypothesis but rather to use the tests. Output scalars are features in the proposed epilepsy detection and classification. Overall, five different hypotheses/statistical tests are implemented in the proposed approach, each of which measures different characteristics of the input data. *h* and *p*-values, which are extracted from each test, and other scalars are calculated and used as features in the proposed approach. The set of parameters that are extracted and used for epilepsy detection and classification are listed in Table 1.

Table 1. Set of parameters in proposed approach.

Test	Parameters
Chi-Square	*h, p, degree of freedom*
Durbin–Watson	*h, p*
Run test	$h, p, R, n^{+}, n^{-}, Std(R)$
Kruskal–Wallis	$h, p, S_1, S_2, S_{total}, Df_1, Df_2, MSE_1, MSE_2$
z test	*h, p, upper confidence, lower confidence, z value*

3.4. Step 4: Machine-Learning Methods

3.4.1. Dimensionality Reduction and Feature Selection

Feature selection is the process of reducing the number of features in the input set. In this context, two concepts are defined: feature selection and dimensionality reduction. In feature selection, the output set of features is a subset of the original set; in dimensionality reduction, the output can be a new synthetic feature set. Thus, feature selection is a special case of dimensionality reduction. In the proposed approach, feature selection and dimensionality reduction are implemented using two approaches: manual and automatic selection.

3.4.2. Classification

In the classification step, a model is constructed in the training phase using the training samples after the feature selection process. The model is then used during the testing phase to evaluate the performance of the epilepsy detection and classification approach. Various classification algorithms are implemented in the proposed approach, as listed in Table 2.

Table 2. Set of classification algorithms used in the proposed approach.

No.	Algorithm	Type
1	Logistic model tree (LMT)	Decision tree
2	J48	Decision tree
3	Random forest	Decision tree
4	KNN	Instance-based
5	SVM	Support vector machine
6	Naive Bayesian	Probability-based
8	Feed-forward NN	Neural network

4. Experiment Tests

4.1. Dataset

A dataset for epilepsy cases that is publicly available online with a description about the acquisition conditions was constructed by Bonn University. Each sample in the Bonn University dataset represents an EEG signal that was recorded for a duration of 23.6 s by 128 channel amplifier

systems with a sampling rate of 173.61 Hz and a 12 bit resolution. Each sample was acquired for a subject in a specific state. Accordingly, the data consist of five different categories, A, B, C, D, and E, each containing 100 samples. Sets A and B are for five healthy subjects in a relaxed and awake state with open eyes for Set A and closed eyes for Set B. Sets C, D, and E were obtained from five epilepsy patients in different states. Set C was obtained from epilepsy patients in seizure-free intervals recorded from the hippocampal formation of the nonepileptogenic hemisphere of the brain. Set D was recorded for the same state, but for the epileptogenic hemisphere of the brain. Set E was obtained from epilepsy patients in a seizure state. In total, 500 samples were used in the experiments. For conducting the experiments, Sets A and B were combined to form 200 samples for healthy subjects. Similarly, Sets C and D were combined to form 200 samples of interictal states of epilepsy patients, and Set E has 100 samples of ictal states of epilepsy patients [27].

4.2. Normality Test

Before data are used for evaluation, they are first checked for normality. The normality of the input samples is checked using a Chi-square test with $\alpha = 0.05$. As shown in Table 3 (Feature No. 1), all samples in the dataset were shown to have a value of zero, which indicates that normality was present for the input data.

Table 3. Statistics of parameter values with highlighted features not considered in the classification and detection step.

No.	Features	Value Range	Min	Max	Mean	Std
Chi Square						
1	h	0/1	0	0	0	0
2	p	[0–1]	0.994	0.994	0.994	0
3	*freedom*	[1–∞)	4096	4096	4096	0
Durbin–Watson						
4	h	0/1	0	0	0	0
5	p	[0–1]	0.003	0.366	0.093	0.064
Run–Test						
6	h	0/1	1	1	1	0
7	p	[0–1]	0	0	0	0
8	R	[0–∞)	63	876	38,532	142,716
9	n^+	[1–∞)	1292	2718	2,058,052	174,186
10	n^+	[1–∞)	1379	2805	2,038,948	174,186
11	$Std(R)$	(∞–∞)	−61,757	−36,656	−51,883	4,438
Kruskal–Wallis						
12	h	0/1	0	0	0	0
13	p	[0–1]	0	1	0.48	0.314
14	S_1	[1–∞)	677	4096	2,418,56	1,089,515
15	S_2	[1–∞)	1,119,929	1,755,959	1,409,027	29,217
16	S_{total}	[1–∞)	789,905,055	5,730,814,242	3,396,122,003	151,838,272
17	Df_1	[1–∞)	564,712	4096	2,427,396	1,085,198
18	Df_2	[1–∞)	0	1	0.48	0.314
19	MSE_1	[1–∞)	0	4,939,480,042	2,334,411,002	151,816,665
20	MSE_2	[1–∞)	57,290,636,765	57,308,142,425	5,730,533,006,017	284,775,365
z-test						
21	h	0/1	0	0	0	0
22	p	[0–1]	1	1	1	0
23	*upper confidence*	(∞–∞)	−77,813	56,191	−10,936	26,044
24	*lower confidence*	(∞–∞)	−75,751	67,962	−457	26,973
25	*z-value*	(∞–∞)	0	0	0	0

4.3. Feature Selection

In the first step of feature selection, features with low or no variation, as determined by the standard deviation, were removed from the dataset. Accordingly, ten features were removed, as shown in Table 3 (i.e., highlighted features were removed), and only 15 features, representing four tests,

were left after the first feature selection step. Thus, the Chi-square variance test was only used for normalization and was not included in the epilepsy classification. In the second step of feature selection, dimensions were reduced using Principal Component Analysis (PCA). PCA generated five new features out of 15 features that were left after the first feature selection step. Statistics on these features are given in Table 4.

For the experiment, MATLAB R2015a [28] was used. For the hardware part, a 64–bit Windows 10 laptop with Intel Core i7 Processor (2.60 GH), with DDR4 RAM is used.

Table 4. Statistics of features generated in the Principal Component Analysis (PCA).

No.	Min	Max	Mean	Std
1	−3.849	4.939	0	2.202
2	−5.213	4.449	0	1.805
3	−4.37	4.974	0	1.524
4	−5.738	5.254	0	1.443
5	−4.145	4.592	0	1.359

5. Results, Discussion, and Comparisons

In the result evaluation step, the classification process was implemented for two tasks: detection and classification. In each of these tasks, the training and testing processes were implemented n-folds, where n was set to 10. At the n-th fold, the data were equally divided by n fold and experiments were conducted for n rounds. In each round, $n − 1$ fold were used for training and 1-fold were used for testing. Accordingly, each fold was used in the testing of a round so all available data were used. The results are reported as the results of all time points.

Only two classes are presented for the detection task. A comparison between the used classification algorithms prior to and after PCA is given in Figure 4. As noted, the implementation of PCA enhanced the results of the best performed algorithms—Random Forest and K-Nearest Neighborhood—in terms of their performance, while the results of the Neural Network significantly decreased. PCA also slightly decreased the results of J48 and SVM and enhanced the results of the Logistic Model Tree. Finally, the results of the Naïve Bayesian algorithm remained as they were before and after PCA. According to the given results, the proposed approach with PCA dimension reduction achieved an accuracy rate of 98.4%. The Random Forest and K-Nearest Neighborhood algorithms achieved the highest accuracy levels. The Logistic Model Tree, J48 Tree, and Neural Network achieved results that were close to the best results with 97.8%, 97.6%, and 97.2%, respectively. The Naïve Bayesian algorithm achieved an accuracy of 96.8%, and SVM achieved an accuracy of 94.4%. Overall, the proposed approach achieved results with an accuracy of 98.4%. In the classification task, three sample classes are presented, and the relative measures shown are an indication of class prediction against the rest of the classes. Another comparison between the algorithms used prior to and after PCA for the classification task is given in Figure 5.
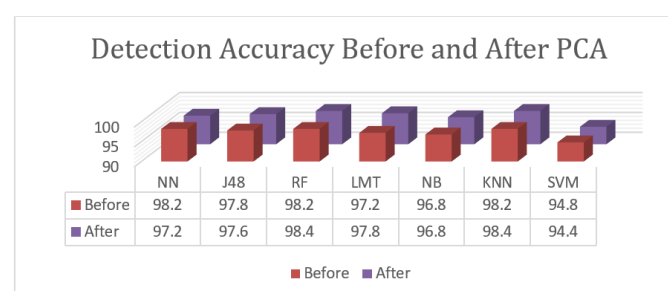

Detection Accuracy Before and After PCA

	NN	J48	RF	LMT	NB	KNN	SVM
Before	98.2	97.8	98.2	97.2	96.8	98.2	94.8
After	97.2	97.6	98.4	97.8	96.8	98.4	94.4

Figure 4. Accuracy comparison between classification algorithms for the detection task before and after the PCA.
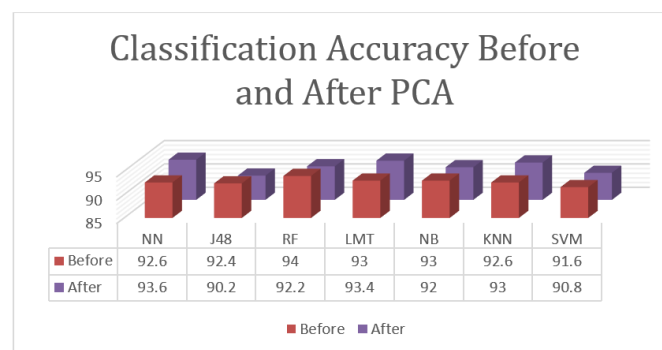
Figure 5. Accuracy comparison between classification algorithms for the classification task before and after the PCA.

According to the given results, the proposed approach for the classification tasks of the three involved states with PCA achieved an accuracy rate of 93.6%, which was reported using the Neural Network. The Logistic Model Tree and K-Nearest Neighborhood algorithms achieved results that were close to the top results with 93.4% and 93.0%, respectively. Random Forest achieved an accuracy of 92.2%, Naïve Bayesian achieved an accuracy of 92.0%, SVM achieved an accuracy of 90.8%, and J48 achieved and accuracy of 90.2%. While PCA enhanced the results of some algorithms, it failed to enhance the high accuracy and decreased the best results from 94.0% to 93.6%. PCA enhanced the results of the Neural Network, Logistic Model Tree, and K-Nearest Neighborhood algorithms but decreased the results of the other algorithms.

Overall, the proposed approach achieved an accuracy of 94.0% for the classification task, which is lower than the accuracy of the detection because this task is more complicated than detection as it involves three classes instead of two.

As noted, the proposed approach achieved high results with an accuracy of 94.0% and 98.4% for classification and detection. This indicates that the proposed approach has the ability to be reliably used in the classification of epileptic seizures. The Random Forest algorithm is the best to be used with PCA in the detection and without PCA in the classification. The SVM algorithm gave the worst results overall.

5.1. Comparison

This section provides a comparison between the proposed approach and state-of-the-art methods defined in the literature. The comparison was done on the basis of both the classification and detection tasks. The accuracy and time complexity of the proposed method are not discussed in detail.

Table 5 summarizes the results of the proposed and the compared approaches for the detection task. As noted, the proposed approach overperformed when compared with the state-of-the-art methods for epilepsy detection on the Bonn dataset. The rest of the literature could not be compared because the results were computed based on real-data or using only the undetermined part of the Bonn dataset. Table 6 summarizes the results of the proposed and compared approaches for the classification task. As noted, the proposed approach was fifth out of six; while it ranked low, it still outperformed one of the recent approaches for epilepsy classification.
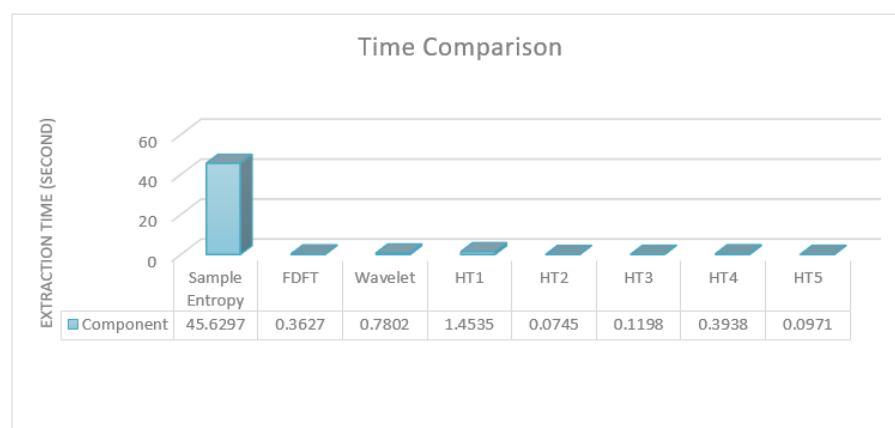
Table 5. Comparison between the proposed approach and state-of-the-art methods for the detection task.

Ref.	Features	Classifier	Data	Classes	Acc.
Proposed	Hypothesis Features	Random Forest	Bonn	Ictal vs. others	98%
Guo et al. [23]	Wavelet and line-length	ANN	Bonn	Ictal vs. others	97.77%

Table 6. Comparison between the proposed approach and state-of-the-art methods for the classification task.

Ref.	Features	Classification	Dataset	Results
S. Vijith et al. [9]	Approximate entropy	SVM	Bonn	89%, 91%
The proposed approach	Hypothesis Features	Random Forest	Bonn	94.0%
Orhan et al. [24]	Wavelet coefficients	ANN	Bonn	95.6%
Acharya et al. [5]	Non–linear	SVM and GMM	Bonn	96.1%
Krishnakumar and Thanushkodi [13]	STFT and non-linear	ANN	Bonn	96.2%
Acharya et al. [7]	Non-linear	SVM, KNN, FC, ANN, DT, GMM, and NBC	Bonn	98.1%

The idea was to compare the performance of the proposed approach against the state-of-the-art ones defined in the literature from different perspectives. One of the samples in the Bonn dataset was used as input to various processing and feature extraction techniques, which are reported in the literature. The time taken to extract entropies, which are the commonly used features, and carry out STFT, wavelet, and each of the hypothesis tests is reported. Figure 6 illustrates the time taken for the compared and proposed feature extraction.

Figure 6. Time comparison between the proposed feature extraction method versus state-of-the-art feature extraction methods.

As noted, all of the used tests, with the exception of the Chi-square variance, which was eliminated in the first feature selection process, consumed less time than the others. Indeed, the four tests that were used in the proposed approach consumed less time compared to wavelet transformation alone, which should usually be followed by the feature extraction step. Sample entropy, which is accurate for epilepsy detection and classification, took over 100 times longer than the worst of the four used tests in the proposed approach. Please note that most of the compared approaches use more than one feature as well as using transformations with features, which significantly increases their time requirement. Thus, in terms of implementing a real-time analysis with reliable results, the proposed approach could be one of the best choices due to its low time consumption and high accuracy. Overall, the proposed approach had very accurate results in the detection task and outperformed the state-of-the-art methods in a similar task on the same dataset. The results also showed that the proposed approach had accurate results in the classification task; however, it did not outperform the state-of-the-art methods even though it consumed less time.

6. Conclusions

In this paper, a new approach that enables the efficient, accurate, and nonparametric detection and classification of epileptic seizures based on EEG signals was proposed. The proposed approach has two components: feature extraction and the use of a classification algorithm. Regarding the features, hypothesis test results were used as input features. Accordingly, PDF was estimated from a particular observed data sample. Then, hypothesis tests of various types were implemented. The purpose of the implemented test depends on the type of the test; this includes quality data fitting with their

PDF as well as data randomness and correlation. The proposed approach achieved high results with accuracy levels of 94.0% and 98.4% for classification and detection, respectively. This indicates that the proposed approach could reliably be used in the classification of epileptic seizure cases. All of the used tests, with the exception of the Chi-square variance, which was eliminated in the first feature-selection process, consumed less time than other approaches in the literature.

Author Contributions: This work is part of a Master's thesis submitted by A.A. for the fulfilment of the Master's degree in Computer Science at Mutah University, Jordan. The idea of the work was conceptualized by A.H. Methodology and validation were provided by R.A. Software implementation and visualization were done by A.A. Writing—original-draft preparation was done by A.A., and writing—review and editing were done by R.A. Project administration was done by A.H. and M.A.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fisher, R.S.; van Emde Boas, W.; Blume, W.; Elger, C.; Genton, P.; Lee, P.; Jerome Engel, J. Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia* **2005**, *46*, 470–472. [CrossRef] [PubMed]
2. WHO. Epilepsy: A Public Health Imperative: Summary. 2019 . Available online: https://www.who.int/mental_health/neurology/epilepsy/report_2019/en/ (accessed on 1 July 2019)
3. Van Mierlo, P.; Papadopoulou, M.; Carrette, E.; Boon, P.; Vandenberghe, S.; Vonck, K.; Marinazzo, D. Functional brain connectivity from EEG in epilepsy: Seizure prediction and epileptogenic focus localization. *Prog. Neurobiol.* **2014**, *121*, 19–35. [CrossRef]
4. Coito, A.; Genetti, M.; Pittau, F.; Iannotti, G.; Thomschewski, A.; Hller, Y.; Trinka, E.; Wiest, R.; Seeck, M.; Michel, C.; et al. Altered directed functional connectivity in temporal lobe epilepsy in the absence of interictal spikes: A high density EEG study. *Epilepsia* **2016**. [CrossRef] [PubMed]
5. Acharya, U.R.; Chua, C.K.; Lim, T.C.; Dorithy, Suri, J.S. Automatic identification of epileptic EEG signals using nonlinear parameters. *J. Mech. Med. Biol.* **2009**, *9*, 539–553. [CrossRef]
6. Kannathal, N.; Choo, M.L.; Acharya, U.R.; Sadasivan, P.K. Entropies for Detection of Epilepsy in EEG. *Comput. Methods Prog. Biomed.* **2005**, *80*, 187–194. [CrossRef] [PubMed]
7. Acharya, U.R.; Molinari, F.; Sree, S.V.; Chattopadhyay, S.; Ng, K.H.; Suri, J.S. Automated diagnosis of epileptic EEG using entropies. *Biomed. Signal Process. Control.* **2012**, *7*, 401–408. [CrossRef]
8. Yang, Z.; Wang, Y.; Gaoxiang, O. Adaptive Neuro-Fuzzy Inference System for Classification of Background EEG Signals from ESES Patients and Controls. *Sci. World J.* **2014**, *2014*, 140863. [CrossRef] [PubMed]
9. Vijith, V.S.; Jacob, J.E.; Iype, T.; Gopakumar, K.; Yohannan, D.G. Epileptic seizure detection using non linear analysis of EEG. In Proceedings of the 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 26–27 August 2016; pp. 1–6. [CrossRef]
10. Thilagaraj, M.; Pallikonda Rajasekaran, M.; Arun Kumar, N. Tsallis entropy: As a new single feature with the least computation time for classification of epileptic seizures. *Clust. Comput.* **2018**. [CrossRef]
11. Li, P.; Karmakar, C.; Yearwood, J.; Venkatesh, S.; Palaniswami, M.; Liu, C. Detection of epileptic seizure based on entropy analysis of short-term EEG. *PLoS ONE* **2018**, *13*, e0193691. [CrossRef] [PubMed]
12. Nijsen, T.M.; Cluitmans, P.J.; Griep, P.A.; Aarts, R.M. Short Time Fourier and Wavelet Transform for Accelerometric Detection of Myoclonic Seizures. In Proceedings of the 1st IEEE/EMBS Benelux Symposium, Brussels, Belgium, 7–8 December 2006.
13. Krishnakumar, S.; Thanushkodi, K. An improved EEG signal classification using Neural Network with the consequence of ICA and STFT. *J. Electr. Eng. Technol.* **2014**, *9*, 1060–1071. [CrossRef]
14. Kovcs, P.; Samiee, K.; Gabbouj, M. On application of rational Discrete Short Time Fourier Transform in epileptic seizure classification. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; pp. 5839–5843. [CrossRef]
15. Samiee, K.; Kovcs, P.; Gabbouj, M. Epileptic Seizure Classification of EEG Time-Series Using Rational Discrete Short–Time Fourier Transform. *IEEE Trans. Biomed. Eng.* **2014**, *62*, 541–552. [CrossRef] [PubMed]

16. Szuflitowska, B.; Orlowski, P. Comparison of the EEG Signal Classifiers LDA, NBC and GNBC Based on Time-Frequency Features. *Pomiary Autom. Robot.* **2017**, *21*, 39–45. [CrossRef]

17. Übeyli, E.D.; Cvetkovic, D.; Holland, G.; Cosic, I. Adaptive Neuro-fuzzy Inference System Employing Wavelet Coefficients for Detection of Alterations in Sleep EEG Activity During Hypopnoea Episodes. *Digit. Signal Process.* **2010**, *20*, 678–691. [CrossRef]

18. Sadati, N.; Mohseni, H.; Maghsoudi, A. Epileptic Seizure Detection Using Neural Fuzzy Networks. In Proceedings of the 2006 IEEE International Conference on Fuzzy Systems, Vancouver, BC, Canada, 16–21 July 2006; pp. 596–600. [CrossRef]

19. Jahankhani, P.; Kodogiannis, V.; Revett, K. EEG Signal Classification Using Wavelet Feature Extraction and Neural Networks. In Proceedings of the IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA'06), Sofia, Bulgaria, 3–6 October 2006; pp. 120–124. [CrossRef]

20. Subasi, A. Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction. *Comput. Biol. Med.* **2007**, *37*, 22–44. [CrossRef] [PubMed]

21. Subasi, A. EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Syst. Appl.* **2007**, *32*, 1084–1093. [CrossRef]

22. Costa, R.P.; Oliveira, P.; Rodrigues, G.; Leitao, B.; Dourado, A. Epileptic Seizure Classification Using Neural Networks with 14 Features. In Proceedings of the Knowledge-Based Intelligent Information and Engineering Systems, Zagreb, Croatia, 3–5 September 2008; Springer: Berlin/Heidelberg, Germany, 2008; pp. 281–288.

23. Guo, L.; Rivero, D.; Dorado, J.; Rabual, J.; Pazos, A. Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. *J. Neurosci. Methods* **2010**, *191*, 101–109. [CrossRef] [PubMed]

24. Orhan, U.; Hekim, M.; Ozer, M. EEG signals classification using the K means clustering and a multilayer perceptron neural network model. *Expert Syst. Appl.* **2011**, *38*, 13475–13481. [CrossRef]

25. Khan, Y.; Farooq, O.; Sharma, P. Automatic Detection of Seizure ONSET in Pediatric EEG. *IJESA* **2012**, *2*, 81–89. [CrossRef]

26. Rioul, O.; Vetterli, M. Wavelets and signal processing. *IEEE Signal Process. Mag.* **1991**, *8*, 14–38. [CrossRef]

27. Winterhalder, M.; Maiwald, T.; Voss, H.; Aschenbrenner-Scheibe, R.; Timmer, J.; Schulze-Bonhage, A. The seizure prediction characteristic: A general framework to assess and compare seizure prediction methods. *Epilepsy Behav. EB* **2003**, *4*, 318–325. [CrossRef]

28. *MATLAB*, Version 8.5 (R2015a); The MathWorks Inc.: Natick, MA, USA, 2015.