Health electroencephalogram epileptic classification based on Hilbert probability similarity

Abdulkareem A. Al-Hamzawi, Dhiah Al-Shammary, Alaa Hussein Hammadi
College of Computer Science and Information Technology, University of Al-Qadisiyah, Dewaniyah, Iraq

Article Info

Article history:

Received Aug 19, 2022
Revised Sep 7, 2022
Accepted Oct 1, 2022

Keywords:

Electroencephalogram classification
Epileptic
Hilbert similarity
Probability similarity
Seizure detection

ABSTRACT

This paper has proposed a new classification method based on Hilbert probability similarity to detect epileptic seizures from electroencephalogram (EEG) signals. Hilbert similarity probability-based measure is exploited to measure the similarity between signals. The proposed system consisted of models based on Hilbert probability similarity (HPS) to predict the state for the specific EEG signal. Particle swarm optimization (PSO) has been employed for feature selection and extraction. Furthermore, the used dataset in this study is Bonn University's publicly available EEG dataset. Several metrics are calculated to assess the performance of the suggested systems such as accuracy, precision, recall, and F1-score. The experimental results show that the suggested model is an effective tool for classifying EEG signals, with an accuracy of up to 100% for two-class status.

Corresponding Author:

Abdulkareem A. Al-Hamzawi
College of Computer Science and Information Technology, University of Al-Qadisiyah
Dewaniyah, Iraq
Email: abdelkarim.karmul@qu.edu.iq

1. INTRODUCTION

An electroencephalogram (EEG) is a recording of the brain's electric activity that can reveal information about brain conditions such as epilepsy and eye conditions. The most common neurological illness in humans is epilepsy, which is characterized by recurring seizures [1], [2]. Seizures are rapid changes in the electrical activity of the brain that cause changed behaviors such as loss of consciousness, jerky movements, temporary lack of breath, and memory loss [3].

Initially, EEG signal processing was purely visual, difficult, time-consuming, and required the assistance of a physician [4]. Changing this outdated classification system proves the difficulty of the process which requires a lot of time and effort. Several studies in biomedical signal processing have focused on the development of classification systems for automatic analysis [5]. Moreover, they concentrated on EEG signals in the detection of epileptic seizures and the identification of eye states EEG [6]. In fact, this can then be integrated into implantable devices that detect the beginning of seizures and trigger a focal treatment to stop or slow the progression of seizures and improve patients' living conditions [7].

EEG signals produce a vast amount of data and visual evaluation of this data achieved by specialists or neurologists is time-consuming and prone to inaccuracy [8]. Consequently, various EEG analysis approaches to computerize the processing of EEG data have been developed. In medical applications, such as epileptic treatment, developers have concentrated on evaluating the most important forms of EEG signals [9]. This paper tackles the issue of detecting epileptic seizures by using another technique for EEG signals classification depending on Hilbert probability similarity (HPS). HPS classifier can detect epileptic seizures from EEG signals within a reasonable time-consuming.

Many medical and service applications rely on the classification of health data [10]. The provision of broad patient monitoring and early diagnosis is essential. Machine learning (ML) algorithms rely heavily on classification methods. On the other hand, they are prone to stagnation, trapped with local optimums, have large time requirements, and produce inconsistent results [11], [12]. In order to overcome traditional classification challenges and drawbacks, potential classification models are technically required. EEG epileptic signal classification is a useful technique for early diagnosis and monitoring of epilepsy seizures.

A novel technique has been developed that exploits the use of HPS to diagnose epileptic seizures using EEG signals. This paper has proposed two models for epileptic seizure detection: the first model is created based on HPS classifier, and the results are significant with achieved accuracy of up to 100% in some classification cases. A second model is built by combining HPS classifier with particle swarm optimization (PSO) as an optimization method. The proposed model HPS with PSO has outperformed the first model in several classification cases according to the accuracy metric results.

In this paper, HPS model has been designed for EEG signal analysis and epileptic seizure detection. Bonn EEG dataset has been utilized for experiment and evaluation. Technically, the proposed model has been compared with other most significant ML algorithms in terms of performance metrics accuracy, precision, Recall, and F1-Score. Another comparison is made with other related work that adopts the same EEG dataset. The experiments are achieved with several EEG dataset testing sizes with the whole EEG signals. Other results are obtained based on various signal lengths. All testing cases are performed with and without optimization. Empirically, the proposed model has clearly outperformed most machine learning algorithms in the same environment.

The rest of this paper is structured as follows. In section 2, some of the related works are explained. Section 3 demonstrates the proposed system with its main concepts. System results are calculated and compared to other EEG classification methods are explained in section 4. Finally, section 5 describes the conclusions and future work.

2. RELATED WORK

EEG classification is essential in a number of EEG-based services and applications. ML methods have been frequently employed for the analysis or prediction of epileptic seizures in raw EEG signals. Technically, a large number of studies concentrate on the detection of epileptic seizures within the EEG signals.

Nkengfack *et al.* [13] discussed the detection and identification of seizure or seizure-free states by using EEG signals for epileptic patients. Discrete Legendre transforms (DLT) and discrete Chebyshev transform (DChT) have been suggested for extraction of beta and gamma rhythms of EEG signal that would be fed as an input to the least square support vector machine (LS-SVM) that is used for the classification process. Technically, accuracy, sensitivity, specificity, and area under the curve (AUC) have been calculated in order to evaluate the proposed model. The evaluation has entirely based on the Bonn university EEG dataset. The best-achieved accuracy result started from 88.75% to 100%.

Mandhouj *et al.* [14] discussed EEG signals classification for normal, pre-ictal, and ictal classes in order to help the detection of epileptic seizure onset. Short-time Fourier transform (STFT) has been utilized for extracting useful information from EEG signals and exploited as an input to the convolution neural network (CNN) classification model. Technically, sensitivity, specificity, accuracy, and precision have been computed to assess the performance of the suggested model. The EEG dataset of Bonn University has been employed for evaluation. The best-achieved result of accuracy was an average rate of 98.22%. On the other hand, standard criteria such have not been used for evaluation such as F1-score, Recall, and AUC. Moreover, processing time has not been computed to investigate how the proposed model is practical.

In order to extract and select the most discriminative features of the EEG signal, Abdelhameed and Bayoumi [8] proposed a variational autoencoder (VAE). VAE is a combination of probabilistic graphical and neural network models that are utilized for EEG signal classification. Technically, accuracy, sensitivity, specificity, precision, and F1-score have been computed in order to evaluate the proposed model. The evaluation has been completely based on two datasets, the Bonn University EEG dataset and the Children's Hospital EEG dataset in Boston. The best accuracy result has been obtained for the Bonn EEG dataset starting from 99% to 100%, and for the Children's Hospital, the EEG dataset started from 96.8% to 99.45%.

Samiee *et al.* [15] discussed the detection and classification of epileptic seizures from EEG patient records, and how is the distinction made between epileptic seizures and normal artifacts with a similar time-frequency paradigm. Discrete short-time Fourier transform (DSTFT) has been generalized to extract features from EEG records. Multilayer perceptron (MLP) architecture has been selected for classification tasks from multiple classifiers been experimented with. MLP has been trained with a back-propagation algorithm. Technically, sensitivity, accuracy, and specificity have been computed for the proposed model evaluation. The evaluation has been completely based on the Bonn university EEG dataset. The best accuracy result has been obtained for two classes starting from 94.9% to 99.8%.

3. PROPOSED METHOD

This section illustrates the complete design for the proposed models including the EEG dataset and HPS measurement. Empirically, cosine similarity measurement, convex set in Hilbert space measurement of similarity, and Hilbert probability-based measure of similarity have been implemented in the proposed method in order to classifying EEG signals and epileptic seizure detection. Hilbert's probability-based measure of similarity has outperformed the other measurements of similarity for EEG signals analysis and classification in terms of performance metrics.

3.1. Hilbert probability similarity measurement

In this paper, HPS measurement has been proposed for EEG signal classification. Mathematically, the HPS is used to compute the similarity between two vectors. Similarity can be obtained using (1) which describes the Hilbert probability formula with its parameters,

$$\text{Similarity} = cos^{-1}(P_1^{\frac{1}{2}} . P_2^{\frac{1}{2}} + Q_1^{\frac{1}{2}} . Q_2^{\frac{1}{2}}) \quad (1)$$

where P1 represents the probability of the most frequency value in the first vector. Similar to P1, P2 represents the probability of the most frequency value in the second vector. The complement probability of p1 is (1-P1) has been depicted by Q1 which mean the probability of the rest values within the first vector. Similarly, to that, the complement probability of P2 is (1-P2) has been described by Q2 [16], [17]. Technically, this similarity function has never been used in machine learning algorithms. The idea of developing Hilbert probability-based similarity in a classification method is the core of the paper. Parameters of this function are exploited to represent the most significant features within the EEG signal.

3.2. EEG dataset

In this paper, the Bonn University EEG dataset has been employed for the experiments and results [18]. It is the most extensively used dataset for identifying epileptic seizures within EEG signals containing five sets of one hundred EEG signals. Each set specifies an identical condition for patients as illustrated in Table 1. Each EEG signal has 23.6 second time duration with 4,097 features.

Table 1. Bonn University EEG dataset

Set Name	Samples	Epoch duration	Length of Segment	Sample frequency	Patient stage	Patient situation
A	100	23.6 s	4097	173.61 (Hz)	Eye open (Normal)	Healthy
B	100	23.6 s	4097	173.61 (Hz)	Eye close (Normal)	Healthy
C	100	23.6 s	4097	173.61 (Hz)	Seizure free (Pre-Ictal)	Epileptic
D	100	23.6 s	4097	173.61 (Hz)	Seizure free (Post-Ictal)	Epileptic
E	100	23.6 s	4097	173.61 (Hz)	Seizure activity (Ictal)	Epileptic

3.3. Hilbert probability similarity classifier

The proposed classification model is mainly based on HPS measurement that is utilized for computing the similarity of two vectors. In our system, these vectors refer to the EEG test signal with each train EEG signal. Each vector represents an EEG signal that is depicted by a large number of data point sampling from the EEG signal. Firstly, the EEG dataset is split into a training set and a testing set. Then for each signal in the testing part, the similarity list is computed using Hilbert similarity-probability-based similarity measure between each testing signal and all training EEG signals. Then, the similarity list is sorted in descending order and the first top five items are considered for the predicted class by computing the peak of the histogram of the EEG class label. Evidently, the decision to pick the top five is based on several experiments using one, two, three, four, five, six, seven, eight, and nine top EEG cases that have empirically proven the highest accuracy can be obtained by the five-selection decision. Finally, the accuracy metric is calculated for all testing signals. All steps of the proposed system are shown in Figure 1.

3.4. Hilbert probability similarity with optimization

In this model, PSO is used to decrease the number of selected features in the EEG signal. PSO produces better results, more quickly and more affordably [19]. PSO is a population-based stochastic optimization technique that mimics animal social behavior like flocks of birds or schools of fish [20]. It begins by randomly selecting a population (swarm) of potential solutions (particles) Generations are updated in order to find the optimal solution. PSO first randomized the velocity and weight of all particles depending on specific parameters. The fitness value for each particle is then determined as the global best position which represents the optimal solution and is iteratively updated. Then, the velocity and position equations are used to update particle velocity and position. Next, the complete optimization process would be repeated as a new iteration with the same steps in order to reach the optimal solution [21]. Figure 2 illustrates the design of the proposed model combined with optimization using the PSO algorithm.

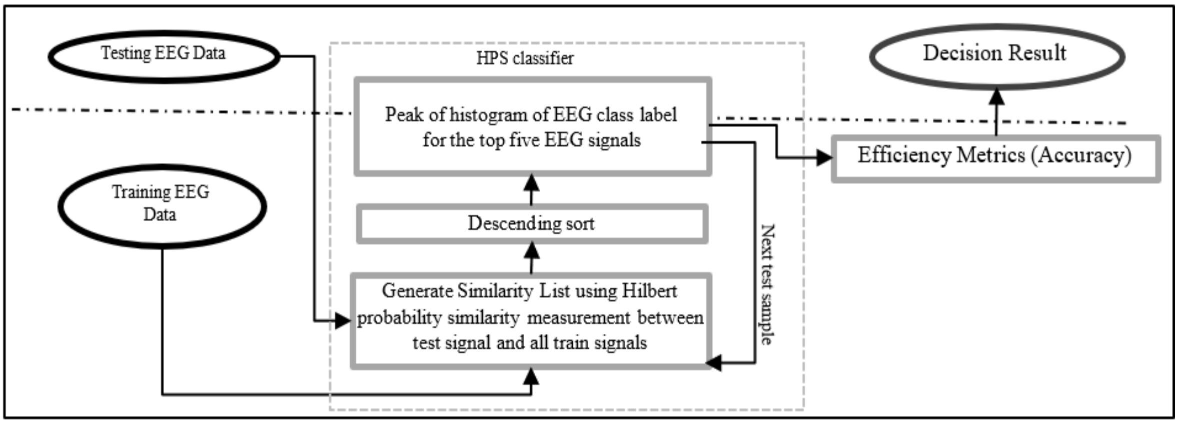

Figure 1. HPS classifier

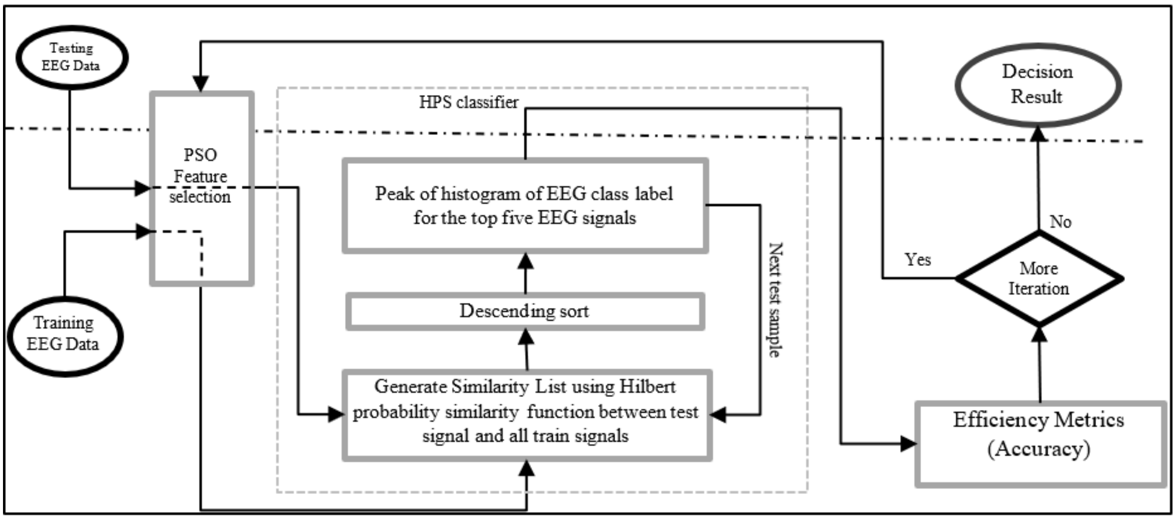

Figure 2. HPS classifier with PSO feature selection

4. EXPERIMENTS AND RESULT

In this section, experiments and results for the proposed system are explained and evaluated based on the suggested models with and without the PSO algorithm. EEG signals of various durations and different testing sizes are performed. Several widely used machine learning algorithms for classification tasks are compared under the same conditions. Other comparisons are achieved with a number of previous studies that are similar to our proposed models using the same dataset. Technically, accuracy, precision, recall, and F1-score have been computed for the experiments and evaluation. Finally, in order to display the results in a comprehensible manner, many illustrative charts have been presented.

4.1. Evaluation strategy

The Bonn University EEG dataset has included five different types of classes. As shown in Table 2, the experiments are concentrated on the class that describes epilepsy seizures. The evaluation and comparison process are carried out across the entire wavelength of EEG signals. The testing results are compared with machine learning classification techniques such as K-nearest neighbor (KNN) with K=3, support vector machine (SVM), random forest (RF), decision tree (DT), and naive Bayes (NB). Finally, another comparison is achieved with previous studies that detailed the same EEG dataset.

Table 2. Evaluation strategy

Test number	EEG signals sets	Description
1	S-Z	Epilepsy seizure and health \ open eyes
2	S-O	Epilepsy seizure and health \ closed eyes
3	S-F	Epilepsy seizure and post-ictal
4	S-N	Epilepsy seizure and pre-ictal
5	S-Z-O	Epilepsy seizure against two cases of healthy people

4.2. Performance metrics

The evaluation of ML algorithms is the most important aspect of any machine learning project. Several measures have been computed for the proposed model evaluation such as accuracy, precision, recall, and F1-score. As a predictive metric, accuracy is the most often used parameter in the classification process. As shown in Figure 3, a confusion matrix is a relationship between the actual class labels and the predicted class labels [22].

Actual Values	Predicted values	
	True Positives (TP) The predicate is true and the actual is true.	False positives (FP) The actual is true and false in the predicate.
	False negatives (FN) The actual is true and false in the predicate.	True negatives (TN) The predicate is false and false in the actual

Figure 3. Confusion matrix

The accuracy measure is the most commonly used performance metric. It is obtained by dividing the total number of calculated results by the number of correct predictions. Accuracy is achieved based on the confusion matrix using (2) [23].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

4.3. Results and analysis

The training set and testing set are obviously the two sets into which machine learning divides the entire dataset. A testing set is used to evaluate the model performance using a number of ML efficiency measures once the proposed model has been built and trained on a training set. Two scenarios have been employed in this experiment, different testing sizes, and various signal lengths.

Scenario 1: Different test sizes

Different testing sizes have been employed in the evaluation strategy in order to properly construct system performance evaluation and also to show how the size of the training set affects the patterns that have developed. Empirically, 10%, 20%, 30%, and 40% of EEG signals testing size are performed during the evaluation process. All computations are carried out without using any preprocessing for the signals with the whole length of the EEG signal (23.6 seconds). Table 3 shows accuracy results for HPS classifier with and without PSO compared to ML algorithms.

The results are obviously more accurate with a larger training size. Technically, this has resulted in a wide range of patterns that assist the system in classifying the different signals of EEG data. Evidently, by applying a 10% testing size without PSO, the proposed model has achieved high rates of accuracy up to 87% in the case of three-class classification and up to 100% in the case of two-class classification. Moreover, the proposed HPS classifier is combined with PSO and the obtained accuracy of two classes has reached up to 100% and 96.7% for three classes of EEG signals. On the other hand, the lowest obtained efficiency is the result of using a 60% training set and a 40% testing set. The lowest obtained accuracy is 91.6 percent for two-class without optimization and 97 percent for two-class with optimization. All resultant accuracy reported in Table 3 is graphically displayed by a bar chart in Figure 4 that depicts the average accuracy for the HPS classifier with and without PSO as well as for various ML techniques for different testing sizes.

Scenario 2: Different signal lengths

EEG signals are having a temporal period of up to 23.6 seconds (available in the Bonn dataset). Our proposed model is tested with different signal lengths to check how the prediction results would vary. The test has included signal lengths of 1, 5, 10, and 15 seconds. Furthermore, the electroencephalogram for the whole signal length provided in Table 3 is subjected to a performance test.

In order to evaluate the performance of the Hilbert classifier based on the probability-similarity model, the accuracy metric has been calculated and reported in Table 4. All results are compared against the most widely used machine learning algorithms for classification, such as NB, SVM, DT, RF, and KNN.

Table 3. Accuracy results for HPS classifier with and without PSO for different testing size

Testing Type	EEG signal duration	Testing size	HPS classifier	HPS with PSO	NB	SVM	RF	DT	KNN
S-Z	23.6 s	10%	**100**	**100**	100	50.0	100	85.0	65.0
S-O	23.6 s	10%	**95.5**	**100**	95.0	50.0	90.0	90.0	55.0
S-F	23.6 s	10%	**95.5**	**100**	90.0	50.0	95.0	80.0	55.0
S-N	23.6 s	10%	**100**	**100**	95.0	50.0	100	80.0	55.0
S-Z-O	23.6 s	10%	**87.0**	**96.7**	73.3	20.0	70.0	70.0	50.0
S-Z	23.6 s	20%	**100**	**100**	100	45	100	77.5	60
S-O	23.6 s	20%	**92.5**	**100**	97.5	45.0	95.0	72.5	47.5
S-F	23.6 s	20%	**95.0**	**100**	92.5	45.0	97.5	75.0	47.5
S-N	23.6 s	20%	**95.0**	**97.5**	95.0	45.0	95.0	87.5	47.5
S-Z-O	23.6 s	20%	**80.0**	**88.3**	76.7	33.3	73.3	58.3	38.3
S-Z	23.6 s	30%	**100**	**100**	100	45.0	100	76.6	61.6
S-O	23.6 s	30%	**93.3**	**95.0**	98.3	48.3	96.6	76.6	50.0
S-F	23.6 s	30%	**91.6**	**95.0**	95.0	48.3	95.0	81.6	50.0
S-N	23.6 s	30%	**95.0**	**95.0**	96.6	48.3	100	90.0	50.0
S-Z-O	23.6 s	30%	**75.6**	**82.2**	78.8	30.3	75.5	67.7	48.3
S-Z	23.6 s	40%	**100**	**100**	100	48.8	97.5	70.0	58.8
S-O	23.6 s	40%	**95.0**	**97.5**	97.5	48.8	97.5	72.5	51.2
S-F	23.6 s	40%	**96.3**	**98.8**	93.8	48.8	96.3	81.3	51.2
S-N	23.6 s	40%	**96.3**	**98.8**	97.5	48.8	98.8	73.8	51.2
S-Z-O	23.6 s	40%	**80.8**	**86.2**	80.0	29.2	75.8	52.5	42.5

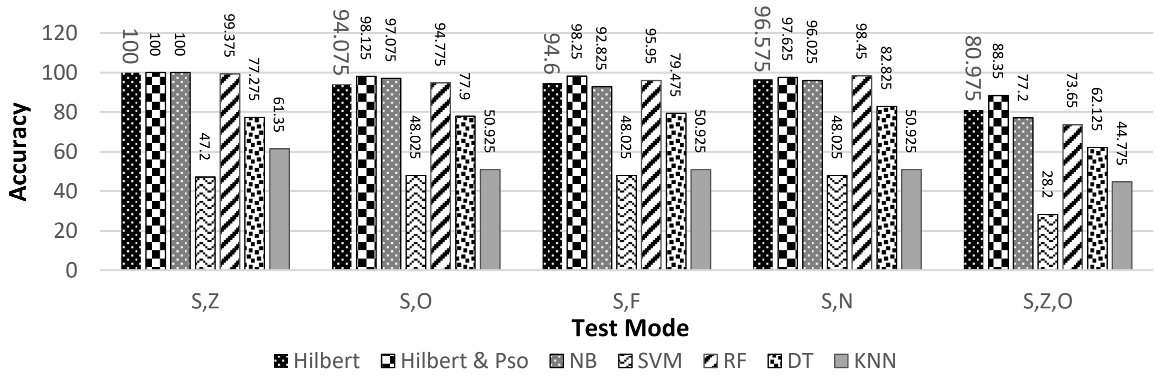

Figure 4. Average accuracy for HPS model with and without PSO compared to machine learning algorithms for different testing sizes

Table 4. Accuracy results for HPS classifier with and without PSO for different signal length

Testing Type	EEG signal duration	Testing size	HPS classifier	HPS with PSO	NB	SVM	RF	DT	KNN
S-Z	1 sec	30%	**91.67**	**93.33**	98.33	48.33	93.33	90.00	81.67
S-O	1 sec	30%	**75.00**	**80.00**	95.00	48.33	90.00	70.00	66.67
S-F	1 sec	30%	**85.00**	**90.00**	86.67	48.33	91.67	83.33	65.00
S-N	1 sec	30%	**83.33**	**86.67**	95.00	48.33	95.00	76.67	63.33
S-Z-O	1 sec	30%	**57.78**	**65.56**	76.67	30.00	77.78	63.33	57.78
S-Z	5 sec	30%	**95.00**	**100.00**	100	48.33	100	76.67	71.67
S-O	5 sec	30%	**83.33**	**90.00**	96.67	48.33	93.33	71.67	55.00
S-F	5 sec	30%	**86.67**	**95.00**	93.33	48.33	93.33	86.67	53.33
S-N	5 sec	30%	**90.00**	**96.67**	95.00	48.33	98.33	81.67	55.00
S-Z-O	5 sec	30%	**76.67**	**80.00**	77.78	30.00	81.11	65.56	46.67
S-Z	10 s	30%	**98.33**	**100**	100	48.33	100	86.67	63.33
S-O	10 s	30%	**83.33**	**93.00**	96.67	48.33	96.67	71.67	53.33
S-F	10 s	30%	**90.00**	**96.00**	93.33	48.33	93.33	81.67	51.67
S-N	10 s	30%	**91.67**	**98.33**	96.67	48.33	100	81.67	53.33
S-Z-O	10 s	30%	**78.89**	**84.44**	78.89	30.00	78.89	53.33	43.33
S-Z	15 s	30%	**100.00**	**100**	100	48.33	100	83.33	61.67
S-O	15 s	30%	**91.67**	**95.00**	96.67	48.33	95.00	68.33	51.67
S-F	15 s	30%	**90.00**	**96.67**	93.33	48.33	91.67	75.00	51.67
S-N	15 s	30%	**91.67**	**96.67**	96.67	48.33	98.33	85.00	51.67
S-Z-O	15 s	30%	**76.67**	**82.22**	78.89	30.00	76.67	60.00	43.33

Clearly, HPS classifier produced a high accuracy of 91.7% without PSO and up to 93.3% with PSO using an EEG signal with a length of one second. With and without PSO, the proposed models have produced better results with an accuracy of up to 100% when the whole EEG signal length (23.6 seconds) has been applied. The minimal accuracy result was reached by using three classes of EEG signals with only 177 features at a one-second signal duration. It has reached 57.8% without PSO and 65.6% with PSO.

The average accuracy metric for 1 sec, 5 sec, 10 sec, 15 sec, and the whole EEG signal length is calculated in order to present the results and performance in such a precise manner for all classification classes. In order to show a clear view of these results Figure 5 depicts the results of the average accuracy metric for different time durations for HPS classifier with and without PSO compared to ML algorithms.

4.4. Comparison with other models

Technically, a high number of alternative techniques have been developed for detecting epileptic seizures. The proposed approach is compared to various previously developed approaches using accuracy measures. This comparison only includes techniques that were implemented within the same dataset, allowing comparisons of outcomes between groups belonging to the same classes.

The results of the comparison in Table 5 show that HPS models have outperformed most of the previous methods. In order to assess the effectiveness of their classifiers, the majority of prior approaches to classifying EEG data have only used two-class classifications. We have categorized various EEG signals and recognized various states of EEG signals as opposed to the methods used by others.

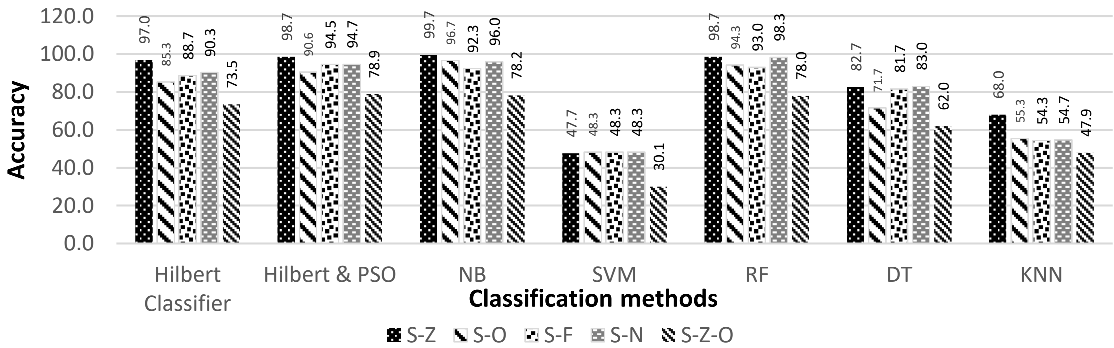

Figure 5. Results of average accuracy metric for HPS classifier with and without PSO of different time durations

Table 5. A comparison of HPS classifier with other previous methods based on best accuracy metric

Author	Methods	Classification case	Best Accuracy
Nkengfack *et al.* [13]	LS-SVM	A – E	100
Samiee *et al.* [15]	DSTFT & MLP	A – E	99.5
Peng *et al.* [24]	Stein kernel-based sparse representation	A – E	99
Liu *et al.* [1]	Energy, ApEn with LPP, LS-SVM	S – Z	98
Attia *et al.* [3]	Burg + SVM	A – E	98
Ech-choudany *et al.* [25]	ANN	A – E	100
Proposed model	**HPS classifier with PSO**	**S - Z**	**100**
Nkengfack *et al.* [13]	LS-SVM	B – E	100
Samiee *et al.* [15]	DSTFT & MLP	B – E	99.3
Peng *et al.* [24]	Stein kernel-based sparse representation	B – E	98.3
Liu *et al.* [1]	Energy, ApEn with LPP, LS-SVM	B – E	----
Attia *et al.* [3]	Burg + SVM	B – E	99
Ech-choudany *et al.* [25]	ANN	B – E	100
Proposed model	**HPS classifier with PSO**	**S – O**	**100**
Nkengfack *et al.* [13]	LS-SVM	C – E	100
Samiee *et al.* [15]	DSTFT & MLP	C – E	98.5
Peng *et al.* [24]	Stein kernel-based sparse representation	C – E	98.3
Liu *et al.* [1]	Energy, ApEn with LPP, LS-SVM	C – E	99.5
Attia *et al.* [3]	Burg + SVM	C – E	99
Ech-choudany *et al.* [25]	ANN	C – E	100
Proposed model	**HPS classifier with PSO**	**S – N**	**100**
Nkengfack *et al.* [13]	LS-SVM	D – E	100
Samiee *et al.* [15]	DSTFT & MLP	D – E	94.9
Peng *et al.* [24]	Stein kernel-based sparse representation	D – E	96.7
Liu *et al.* [1]	Energy, ApEn with LPP, LS-SVM	D – E	98
Attia *et al.* [3]	Burg + SVM	D – E	95
Ech-choudany *et al.* [25]	ANN	D – E	99.5
Proposed model	**HPS classifier with PSO**	**S – F**	**100**

4.5. Time consuming

The proposed classification models (Hilbert classifier based on probability similarity) have outperformed most ML methods for several classification cases within the same environment that is been used for implementation. Python programming language within Windows 10 operating system has been utilized for designing and implementing our models. A personal computer running the system has a 2.30 GHz Intel Core i7-11800H processor and 16 gigabytes of RAM.

The suggested model has an average processing time of 0.05447 ms, a maximum execution time of 0.07720 ms, and a minimum execution time of 0.03505 ms without PSO feature selection. HPS classifier with PSO has an average processing time of 0.02535 ms, minimum execution time of 0.01139 ms, and maximum execution time of 0.054212 ms.

5. CONCLUSION

In this paper, an automated approach for classifying and detecting epileptic seizures from EEG signals is proposed based on HPS as a classifier model and PSO as a feature selection. The Bonn University EEG dataset has been employed for experiments and results. Several performance metrics have been computed for the evaluation such as precision, accuracy, recall, and F1-score. The proposed method is capable of achieving high classification accuracy for two classes reached 100%. In order to properly handle more than two classes, more model development would be performed for future work like adding signals clustering preceding HPS classification.

REFERENCES

[1] Y. Liu, B. Jiang, J. Feng, J. Hu, and H. Zhang, “Classification of EEG signals for epileptic seizures using feature dimension reduction algorithm based on LPP,” *Multimedia Tools and Applications*, vol. 80, no. 20, pp. 30261–30282, Aug. 2021, doi: 10.1007/s11042-020-09135-7.

[2] M. Tosun and Ö. Kasım, “Novel eye-blink artefact detection algorithm from raw EEG signals using FCN-based semantic segmentation method,” *IET Signal Processing*, vol. 14, no. 8, pp. 489–494, Oct. 2020, doi: 10.1049/iet-spr.2019.0602.

[3] A. Attia, A. Moussaoui, and Y. Chahir, “Epileptic seizures identification with autoregressive model and firefly optimization based classification,” *Evolving Systems*, vol. 12, no. 3, pp. 827–836, Sep. 2021, doi: 10.1007/s12530-019-09319-z.

[4] Ö. Kasim and M. Tosun, “Biometric authentication from photic stimulated EEG records,” *Applied Artificial Intelligence*, vol. 35, no. 15, pp. 1407–1419, Dec. 2021, doi: 10.1080/08839514.2021.1981660.

[5] A. A. Alsakaa, M. H. Hussein, Z. H. Nasralla, H. Alsaqaa, K. Nermend, and A. Borawska, “Effective electroencephalogram based epileptic seizure detection using support vector machine and statistical moment’s features,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 5, pp. 5204–5213, Oct. 2022, doi: 10.11591/ijece.v12i5.pp5204-5213.

[6] J. Rabcan, V. Levashenko, E. Zaitseva, and M. Kvassay, “EEG signal classification based on fuzzy classifiers,” *IEEE Transactions on Industrial Informatics*, vol. 18, no. 2, pp. 757–766, Feb. 2022, doi: 10.1109/TII.2021.3084352.

[7] Y. Jiang, W. Chen, M. Li, T. Zhang, and Y. You, “Synchroextracting chirplet transform-based epileptic seizures detection using EEG,” *Biomedical Signal Processing and Control*, vol. 68, Jul. 2021, doi: 10.1016/j.bspc.2021.102699.

[8] A. M. Abdelhameed and M. Bayoumi, “Semi-supervised EEG signals classification system for epileptic seizure detection,” *IEEE Signal Processing Letters*, vol. 26, no. 12, pp. 1922–1926, Dec. 2019, doi: 10.1109/LSP.2019.2953870.

[9] M. Radman, M. Moradi, A. Chaibakhsh, M. Kordestani, and M. Saif, “Multi-feature fusion approach for epileptic seizure detection from EEG signals,” *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3533–3543, Feb. 2021, doi: 10.1109/JSEN.2020.3026032.

[10] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, and C. Fookes, “Deep learning for patient-independent epileptic seizure prediction using scalp EEG signals,” *IEEE Sensors Journal*, vol. 21, no. 7, pp. 9377–9388, Apr. 2021, doi: 10.1109/JSEN.2021.3057076.

[11] X. Lu, J. Zhang, S. Huang, J. Lu, M. Ye, and M. Wang, “Detection and classification of epileptic EEG signals by the methods of nonlinear dynamics,” *Chaos, Solitons & Fractals*, vol. 151, Oct. 2021, doi: 10.1016/j.chaos.2021.111032.

[12] A. A. Al-Hamzawi, D. Al-Shammary, and A. H. Hammadi, “A survey on healthcare EEG classification-based ML methods,” in *Mobile Computing and Sustainable Informatics*, 2022, pp. 923–936, doi: 10.1007/978-981-19-2069-1_64.

[13] L. C. Djoufack Nkengfack, D. Tchiotsop, R. Atangana, V. Louis-Door, and D. Wolf, “Classification of EEG signals for epileptic seizures detection and eye states identification using Jacobi polynomial transforms-based measures of complexity and least-square support vector machine,” *Informatics in Medicine Unlocked*, vol. 23, 2021, doi: 10.1016/j.imu.2021.100536.

[14] B. Mandhouj, M. A. Cherni, and M. Sayadi, “An automated classification of EEG signals based on spectrogram and CNN for epilepsy diagnosis,” *Analog Integrated Circuits and Signal Processing*, vol. 108, no. 1, pp. 101–110, Jul. 2021, doi: 10.1007/s10470-021-01805-2.

[15] K. Samiee, P. Kovacs, and M. Gabbouj, “Epileptic seizure classification of EEG time-series using rational discrete short-time Fourier transform,” *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 541–552, Feb. 2015, doi: 10.1109/TBME.2014.2360101.

[16] Y. C. Sagala, S. Hariyanto, Y. D. Sumanto, and T. Udjiani, “The distance between two convex sets in Hilbert space,” in *AIP Conference Proceedings*, 2021, doi: 10.1063/5.0041684.

[17] W. K. Wootters, “Statistical distance and Hilbert space,” *Physical Review D*, vol. 23, no. 2, pp. 357–362, Jan. 1981, doi: 10.1103/PhysRevD.23.357.

[18] S. S. Al-Fraiji and D. Al-Shammary, “EEG signals classification based on mathematical selection and cosine similarity,” *Journal of Al-Qadisiyah for Computer Science and Mathematics*, vol. 13, no. 3, pp. 57–67, 2021, doi: 10.29304/jqcm.2021.13.3.837.

[19] T. M. Shin, A. Adam, and A. F. Z. Abidin, “A comparative study of PSO, GSA and SCA in parameters optimization of surface grinding process,” *Bulletin of Electrical Engineering and Informatics*, vol. 8, no. 3, pp. 1117–1127, Sep. 2019, doi: 10.11591/eei.v8i3.1586.

[20] R. Eberhart and J. Kennedy, "New optimizer using particle swarm theory," *Proceedings of the International Symposium on Micro Machine and Human Science*, pp. 39–43, 1995, doi: 10.1109/mhs.1995.494215.
[21] A. P. Yoganandini and G. S. Anitha, "A modified particle swarm optimization algorithm to enhance MPPT in the PV array," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 5, pp. 5001–5008, Oct. 2020, doi: 10.11591/ijece.v10i5.pp5001-5008.
[22] S. S. Al-Fraiji and D. Al-Shammary, "Survey for electroencephalography EEG signal classification approaches," in *Mobile Computing and Sustainable Informatics*, 2022, pp. 199–214, doi: 10.1007/978-981-16-1866-6_14.
[23] S. A. Shams, A. Hekal Omar, A. S. Desuky, M. T. Abou-Kreisha, and G. A. Elsharawy, "Even-odd crossover: a new crossover operator for improving the accuracy of students' performance prediction," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 4, pp. 2292–2302, Aug. 2022, doi: 10.11591/eei.v11i4.3841.
[24] H. Peng *et al.*, "Automatic epileptic seizure detection via Stein kernel-based sparse representation," *Computers in Biology and Medicine*, vol. 132, May 2021, doi: 10.1016/j.compbiomed.2021.104338.
[25] Y. Ech-Choudany, D. Scida, M. Assarar, J. Landré, B. Bellach, and F. Morain-Nicolier, "Dissimilarity-based time–frequency distributions as features for epileptic EEG signal classification," *Biomedical Signal Processing and Control*, vol. 64, Feb. 2021, doi: 10.1016/j.bspc.2020.102268.

BIOGRAPHIES OF AUTHORS

Abdulkareem A. Al-Hamzawi received a B.Sc. degree in computer sciences from the University of Al-Qadisiyah, Iraq in 2005. Currently, he is an M.Sc. student at the College of Computer Science and Information Technology, University of Al-Qadisiyah, Iraq. He is interested in artificial intelligence and machine learning techniques, and he has several publications in these fields. He can be contacted at abdelkarim.karmul@qu.edu.iq.

Dhiah Al-Shammary received his Ph.D. in computer science in 2014 from RMIT University, Melbourne, Australia. Dhiah is awarded as the best Ph.D. student and top publication during his Ph.D. period. He has several years of experience in both education and industry. His main industrial experience came from Silicon Valley-based companies working on security projects including non-traditional and quantum-scale encryption. Dhiah has worked at several universities in both Australia and Iraq like RMIT University and the University of Al-Qadisiyah. His research interests include performance modeling, web services, compression and encoding techniques, and distributed systems. Dhiah has several publications in the areas of improving the performance of web services and encoding techniques. He can be contacted at d.alshammary@qu.edu.iq.

Alaa Hussein Hammadi received his Ph.D. from Udmurt State University in Russia. He is an assistant professor in applied mathematics. He is an academic at the College of Computer Science and Information Technology at the University of Al-Qadisiyah, Iraq. He has several publications in the area of applied mathematics and modeling systems in Information Technology. He can be contacted at alaa.hammadi@qu.edu.iq.