Original Article

Detection of Epileptic Seizures from EEG Signals Using Machine Learning Classifiers

Avijit Dey Joy, Sraboni Sarkar and Abul Kalam Azad*

Department of Electrical and Electronic Engineering, University of Dhaka, Dhaka-1000, Bangladesh

* Corresponding Author email: azad@du.ac.bd

Abstract

Epileptic seizure is a chronic neurological disorder which affects millions of people all over the globe. It can be treated in a better way if the symptoms are detected at an early stage. In this study, we have demonstrated and evaluated the classification performances of different machine learning classifiers for the detection of epileptic seizures from electroencephalography (EEG) signals. For this, we have first applied principal component analysis (PCA) on EEG signals to obtain much reduced-length PCA vectors. These vectors are then applied to decision tree (DT), k-nearest neighbor (KNN), Naïve Bayes (NB), support vector machine (SVM) and artificial neural network (ANN) classifiers for the detection of epileptic seizures. The effects of length of PCA vectors on the performances of these classifiers have also been analyzed rigorously for 2-class, 3-class and 5-class classification of EEG signals. Besides such PCA-based classifiers, we have also proposed and evaluated the performances of a customized convolutional neural network (CNN) to directly extract features from the EEG signals as well as to perform classification tasks. The results showed that CNN outperforms PCA-based machine learning classifiers. For 2-class classification cases, CNN attains classification accuracies in the range from 99.50% to 100%, whereas 98.48% and 96.32% accuracies are obtained with CNN for 3-class and 5-class classification cases. The results signify that the proposed CNN classifier can be considered as a highly-efficient scheme for the reliable detection of epileptic seizures from EEG signals.

Keywords: *Classification, Epileptic seizure, EEG signals, Principal component analysis, Machine learning classifiers, Convolutional neural network.*

Introduction

Epileptic seizure is a chronic brain disorder caused by abrupt anomaly of brain nerve cells. It gives rise to abnormal electrical activity in human brain due to the improper firing of brain neurons [1-5]. Seizures have intricate and distinct traumatic symptoms including imbalance body movement, jerking of body parts, loss of consciousness, stroke and disruption of cognitive functions [2, 4]. Severe occurrence of epileptic seizure may even cause death if the patient during seizure attack is unattended [3, 5]. Around the globe, approximately 50 million people suffer from epileptic seizure; among them nearly 80% patients live in developing countries [2, 3]. As stated by the World Health Organization (WHO), there are nearly 2 million patients with epilepsy in Bangladesh [3]. The detection process of epileptic seizures can often be expensive, inappropriate and time-consuming due to the lack of adequately trained medical staff. In managing epilepsy, the erroneous diagnosis of epileptic seizure often creates a treatment gap, particularly in least developed nations like Bangladesh [5].

Electroencephalography (EEG) is a useful technique for the recognition of epileptic seizures. In recording EEG signals, small sensors are affixed to the scalp for collecting the electrical signals generated due the activities in brain. In practical applications, the collected EEG signals are analyzed visually by the physicians to detect the state of epileptic seizure [6, 7]. Such visual inspection is time-consuming and may introduce diverse opinions on the diagnostic results when inspected by different physicians having different levels of diagnostic experiences [7, 8]. Consequently, it is crucial to develop a reliable, robust and computerized system for the effective detection of epileptic seizures from EEG signals.

To detect epileptic seizures, raw EEG signals can be applied directly to suitable classifiers [9]. However, the robust, time-efficient and automated detection process requires the extraction of suitable features from the EEG signals. These extracted features are then applied to classifiers for the accurate detection of epileptic seizures [6-9]. In the literature, various seizure detection techniques have been proposed and numerous studies have been carried out to improve the performance of epileptic seizure detection from EEG signals. In cases of 2-class classification problems, the maximum classification accuracies for using DT classifiers have been reported to be 90.19% with non-leaf node labeling of decision tree based feature extraction [10], 91.97% with variance threshold feature selection [11], 100% with tertiary wavelet transform based features extraction [12] and 99.44% with wavelet transform based feature extraction [13]. For such 2-class classification problems, KNN can provide accuracies in the range from 93.93% to 100% as reported in the research works [10-16] in which 99.81% and 100% accuracies have been reported in the research works [13,14], respectively with wavelet transform based feature extraction. The classification accuracies of 94.3% [10], 95.73% [11] and 100% [14] have been reported for 2-classing classification using NB classifiers. The classification accuracy using multiscale entropy based SVM [6], non-leaf node labeling of decision tree based feature extraction based SVM [10], wavelet analysis based SVM [12, 15], principal component analysis based SVM [17], weighted permutation entropy based SVM [18], have been maximized to be 92.5%, 94.48%, 100%, 100% and 93.38%, respectively. The use of ANN with wavelet transform [13], weighted permutation entropy [18], PCA [19], Stockwell transform [20] and time-frequency domain features [21] can also offer classification accuracies of 99.07%, 91.875%, 97.55%, 90% and 96.60%, respectively for 2-class classification cases.

Recently, deep learning algorithms based CNN have found widespread applications in the classification and detection of various signals [22-24]. The utmost advantage of using CNN is that it offers the enormous capability to extract features from the signals. In such scheme, raw EEG signals can directly be applied to the CNN for highly-accurate detection of epileptic seizures [22-24]. For instances, the 2-class classification of EEG signals can be accomplished with accuracies of 99.86% [22], 100% [23] and 97.03% using deep learning algorithm based

CNN. For 3-class classification of EEG signals, CNN classifiers can provide accuracies of 94.46% [22] and 97.07% [23], which are much better than 91.70% accuracy obtained with multiscale entropy based SVM [6]. For 5-class classification, the accuracy for using CNN have been reported to be 95.84% in the research work [23], which is also comparable to the classification accuracy of 96% obtained with octal pattern and wavelet transform based KNN [25].

Within this framework, the main objective of this study is to develop optimal model for highly-accurate detection of epileptic seizure from EEG signals. For this, we have first used PCA to obtain reduced-length PCA vectors from EEG signals. Then, these PCA vectors have been fed to different machine learning classifiers such as, DT, NB, KNN, SVM and ANN for the detection of epileptic seizure. The optimal length of PCA vectors to be effectively used by a particular machine learning classifier has been explored systematically. In addition, we have also proposed the use of a customized CNN classifier for the accurate detection of epileptic seizure from EEG signals. The performances of these classifiers have been analyzed rigorously and the optimal classifier for epileptic seizure detection has been identified by evaluating the performances of different classifiers in terms of accuracy, sensitivity, specificity and precision.

Materials and Methods

EEG Dataset

The EEG dataset utilized in this study were recorded by Andrzejak et al. at the University of Bonn, Germany [26]. This publicly available dataset [27] comprises five groups of EEG signals (i.e., A, B, C, D and E). Each group consists of 100 single-channel EEG segments. A brief description of EEG signals in Bonn dataset is listed in Table 1.

Table 1. Outlines of EEG signals in Bonn dataset

Class	Recorded Position	Subject State	Recoded period
A	Cortex	Normal	Awake and calm condition with eye open
B	Cortex	Normal	Awake and calm condition with eye closed
C	Hippocampal formation	Epilepsy	Without seizure intervals
D	Epileptogenic Zone	Epilepsy	Without seizure intervals
E	Epileptogenic Zone	Epilepsy	With Seizure activity

The EEG signals in the Bonn dataset was gathered by using a 128-channel amplifier and electrode placement following the popular 10-20 scheme [26]. The recordings of these signals were accomplished at the sampling rate of 173.61 Hz with 12-bit resolution employing a band-pass filter setting of 0.53 Hz to 40 Hz. Each segment of EEG signal lasts for 23.6 seconds having a total of 4097 samples. No preprocessing is performed on the EEG signals included in the Bonn dataset. For instances, one sample EEG signal from each of the five groups in the Bonn dataset are shown in Figure 1.

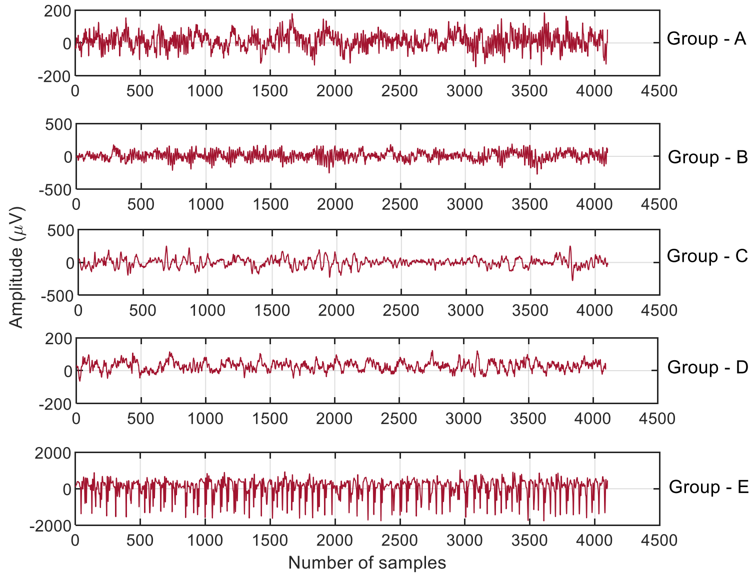

Figure 1. Sample EEG signals from each of the five groups in the Bonn dataset.

In this study, we have explored the performances of a number of machine learning classifiers for 2-class, 3-class and 5-class classification of EEG signals for the reliable detection of epileptic seizures. The summary of the classification cases and classifiers used in this study is listed in Table 2.

Table 2. Classification cases and classifiers used in this study

<table>
<tr><th>Classification</th><th>Class</th><th>Classifiers</th></tr>
<tr><td rowspan="5">2-class</td><td>A-E</td><td rowspan="7">(1) PCA-based DT
(2) PCA-based KNN
(3) PCA-based NB
(4) PCA-based SVM
(5) PCA-based ANN</td></tr>
<tr><td>B-E</td></tr>
<tr><td>C-E</td></tr>
<tr><td>D-E</td></tr>
<tr><td>ABCD-E</td></tr>
<tr><td>3-class</td><td>AB-CD-E</td></tr>
<tr><td>5-class</td><td>A-B-C-D-E</td></tr>
</table>

The classifiers in Table 2 have been realized by MATLAB R2021a on a computer with Intel® Core™ i7-117000 @ 2.50 GHz, 250 GB NVMe SSD, 6GB Graphics card and 16GB DDR4 RAM.

PCA based Machine Learning Classifiers

The Karhunen-Loève transform, commonly known as PCA, is a highly-effective statistical process for reducing the dimensionality of data [9,11,28]. The transformation of data using PCA generates a set of uncorrelated variables, called principal components (PCs). In this study, the PCs of an EEG signal are collectively termed as PCA vector. Each EEG signal in Bonn dataset contains a large number of samples of length P = 4097. For such large P, the computational complexity for using machine learning classifiers will be fairly high. The goal of applying PCA is to attain the reduced-length PCA vectors of EEG signals without significant loss of information so as to minimize the computational complexity of the classifiers.

In the proposed technique of using PCA, the EEG signals to be utilized in the training and testing phases of each of the classifiers form two groups G_{train} and G_{test} comprising EEG signals of length P. Let the numbers of EEG signals in G_{train} and G_{test} groups are Q and R, respectively. If the EEG signals in G_{train} are represented as vectors a_q, where q = 1, 2, 3, …, Q, all the EEG signals in G_{train} can be represented together as a matrix $G_{train} = [a_1, a_2, a_3, \ldots, a_Q]$ of size $P\times Q$ in which, each column of G_{train} consists of a single EEG signal of length P. In a similar fashion, a matrix comprising all EEG signals in G_{test} can be obtained as $G_{test} = [b_1, b_2, b_3, \ldots, b_R]$ of size $P\times R$, where b_r (r = 1, 2, 3,…, R) represent the vector representation of each EEG signal used in the testing phase of classifiers. Now the mean vector $\bar{a}$ for G_{train} is computed as

$$\bar{a} = \frac{1}{Q}\sum_{q=1}^{Q} a_q \qquad (1)$$

Next, the zero-mean matrix $M = [m_1, m_2, \ldots, m_Q]$ are obtained, where $m_q = a_q - \bar{a}$. The covariance matrix V of M can be obtained by

$$V = \frac{1}{Q}\sum_{q=1}^{Q} m_q m_q^T = MM^T \qquad (2)$$

In Eq. (2), the matrix V consists of P eigenvectors in total, also known as PCs and corresponding eigenvalues that are determined by

$$Vk_p = u_p k_p \text{ for } p = 1, 2, 3, \ldots, P \qquad (3)$$

Where, k_p and u_p are the pth eigenvector and eigenvalue of V, respectively. Afterwards, the eigenvectors are ranked in accordance with their eigenvalues. At the end of such process, only L ($L \ll P$) eigenvectors are selected according to the L largest eigenvalues by neglecting all others. The eigenvectors of length L are picked out such that the selection parameter H fulfils the condition given by

$$H = \sum_{p=1}^{L} u_p \Big/ \sum_{p=1}^{P} u_p > t \qquad (4)$$

where, the parameter t is typically selected to be slightly lower than 1 in most practical applications. Now, the vector m in matrix M for a given EEG signal in G_{train} can be estimated as a weighted sum of L selected eigenvectors as given by

$$\sum_{s=1}^{L} c_s k_s \approx m \text{ where } c_s = k_s^T \text{m for } s = 1,2,3,...,L. \qquad (5)$$

A vector x comprising the weights c_s, i.e.,

$$x = \begin{bmatrix} c_1 & c_2 & \dots & c_L \end{bmatrix}^T \qquad (6)$$

can be considered as the PCA vector of a given EEG signal in G_{train}. Consequently, the PCA vectors x_q, for q = 1, 2, 3, …, Q, for all the Q EEG signals in G_{train} can be obtained by using Eq. (6). Thus, each EEG signal of length P in G_{train} can now effectively be replaced by its resultant PCA vector x of much small length L. Similarly, by utilizing the L eigenvectors obtained for G_{train}, PCA vectors y_r (also of length L) for r = 1, 2, …, R can be obtained for all of the R EEG signals in G_{test}.

Next, we have applied machine learning classifiers on the reduced-length PCA vectors corresponding to the EEG signals obtained for G_{train} and G_{test} for the time-efficient detection of epileptic seizure. The PCA vectors for G_{train} along with their associated class labels have been used to train the classifiers in the training phase. In contrast, the feature vectors for G_{test} have been applied to the trained classifiers to estimate the associated class labels of EEG signals in the testing phase. The purpose of using such reduced-length PCA vectors obtained via PCA is to reduce the computational complexity of machine learning classifiers both in training and testing phases [9]. The process of the detection of epileptic seizures using such PCA-based classifiers is depicted in Figure 2.

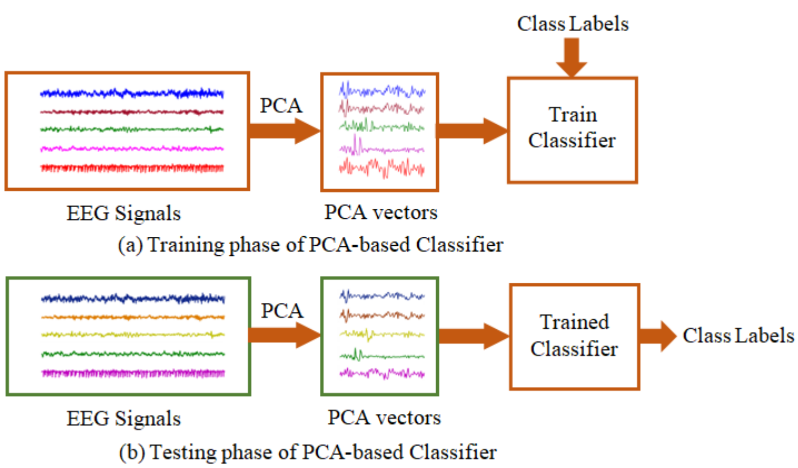

Figure 2. Functional diagram depicting the operating principle of PCA-based classifiers.

As depicted in Figure 2(a), the dimension of EEG signals used in the training phase of machine learning classifiers is first reduced via PCA to obtain reduced-length PCA vectors. These PCA vectors and the associated class labels of corresponding EEG signals are then applied to a particular machine learning classifier to train it. The dimension of EEG signals used in the testing phase is also reduced first via PCA as shown in Figure 2(b). These reduced-length PCA vectors are finally applied to the trained classifier to find the class labels of corresponding EEG signals. In this study, we have adopted a number of widely used machine learning classifiers [9,14], such as DT, KNN, NB, SVM and ANN in the stage of classification for the fast and accurate detection of epileptic seizure.

Convolutional Neural Network (CNN) Classifier

The CNN is a powerful tool that utilizes deep learning algorithm to extract features from signals [22-24]. It also offers the attribute of signal classification. In this study, we propose and demonstrate the use of a customized CNN to directly extract features from the EEG signals as well as to detect epileptic seizures. The layered architecture of the proposed CNN classifier is depicted in Figure 3.

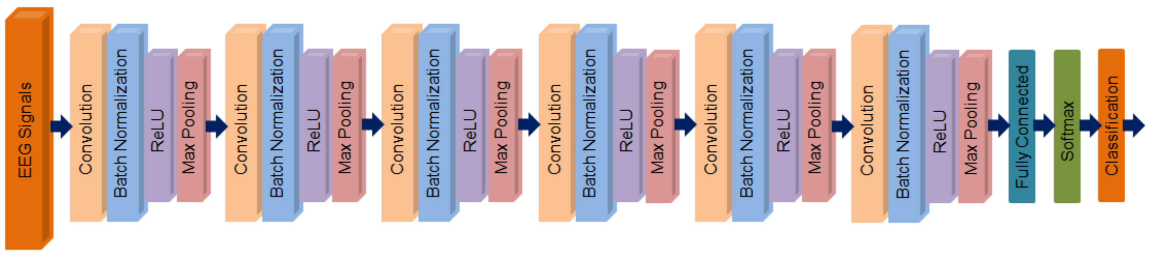

Figure 3. The layered architecture of the CNN classifier.

As shown in Figure 3, the EEG signals are applied to the input layer of the CNN. The input layer is then succeeded by six sets of repetitive and subsequent array of four layers comprising convolutional layer, batch normalization layer, rectified linear unit (ReLU) and Max pooling layer. All of the six convolutional layers used in each set consist of 3×3 filters. The number of filters in the first convolutional layer is fixed to be 16 while that in the second, third, fourth, fifth and sixth convolutional layers are selected to be 32, 64, 128, 256 and 512, respectively. Padding is added in each convolutional layers to make the output size to be equal to the input size. Afterwards, the output of each convolutional layer is normalized by introducing a batch normalization layer. Next, a rectified linear unit (ReLU) is used in each set of layered arrangement to perform the action of nonlinear activation. The use of ReLU also assists to minimize the sensitivity of CNN to the parameter initialization. Finally, a max pooling layer is utilized in each set of layers for the exclusion of redundant information and reduction of size of the extracted features. The feature vectors of the input EEG signals are furnished by the fully connected layer in CNN architecture as shown in Figure 3. The output of the fully connected layer is then normalized by applying the softmax layer. The positive values produced at the output of softmax layer sum to 1. These normalized values are finally used as classification probabilities. Based on these probabilities, the classification layer performs the classification of EEG signals.

Performance Evaluation

In this study, we have adopted 5-fold cross validation to generalize the performances of classifiers. To do so, we have distributed all groups of EEG signals involved in a particular classification case in accordance with Tables 1 and 2. The 5-fold distributions of such EEG signals are depicted in Figure 4.

Portion-01	Portion-02	Portion-03	Portion-04	Portion-05	Fold	Fold Performance (F_i)
Test	Train	Train	Train	Train	Fold-01	F_1
Train	Test	Train	Train	Train	Fold-02	F_2
Train	Train	Test	Train	Train	Fold-03	F_3
Train	Train	Train	Test	Train	Fold-04	F_4
Train	Train	Train	Train	Test	Fold-05	F_5
					Model Performance $= \frac{1}{5}\sum_{i=1}^{5} F_i$	

Figure 4. The distribution of EEG signals and performance evaluation of classifiers in 5-fold cross validation.

As depicted in Figure 4, each of the five folds contains different combinations of 80 EEG signals from each class for the training and remaining 20 EEG signals from each class is used to test the classifier. For a particular classification case, a particular classifier has been used separately for each of these five folds. Then, we have computed four parameters, e.g., *TP* (true positive), *TN* (true negative), *FP* (false positive) and *FN* (false negative) for each class from the confusion matrix obtained for a particular fold. These four parameters are used to compute per-class classification performance in that fold by using Equations (7) - (10).

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \qquad \dots (7)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \qquad \dots (8)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \qquad \dots (9)$$

$$\text{Precision} = \frac{TP}{TP + FP} \qquad \dots (10)$$

The classification performance for a particular fold is then calculated to be the average of the per-class classification performances. Such performances of all of the five folds are eventually averaged to evaluate the performances of the used classifier.

Results and Discussion

In our demonstration, we have analyzed the performances of PCA-based machine learning classifiers as well as that of CNN. For PCA-based machine learning classifiers, we have explored the effect of length (*L*) of PCA vectors obtained for EEG signals. To estimate fixed-length PCA vectors for a particular classifier irrespective of 2-class, 3-class or 5-class classification cases, we have determined the classification accuracies of that classifier by varying the length of PCA vectors from L = 1 to L = 100 at a step of 1. The accuracies provided by the classifier involving all classification cases of A-E, B-E, C-E, D-E, ABCD-E, AB-CD-E and A-B-C-D-E are then averaged to find the overall accuracy of that classifier corresponding to *L*. Such overall accuracies for using DT, KNN, NB, SVM and ANN with PCA vectors having different lengths are shown in Figure 5.

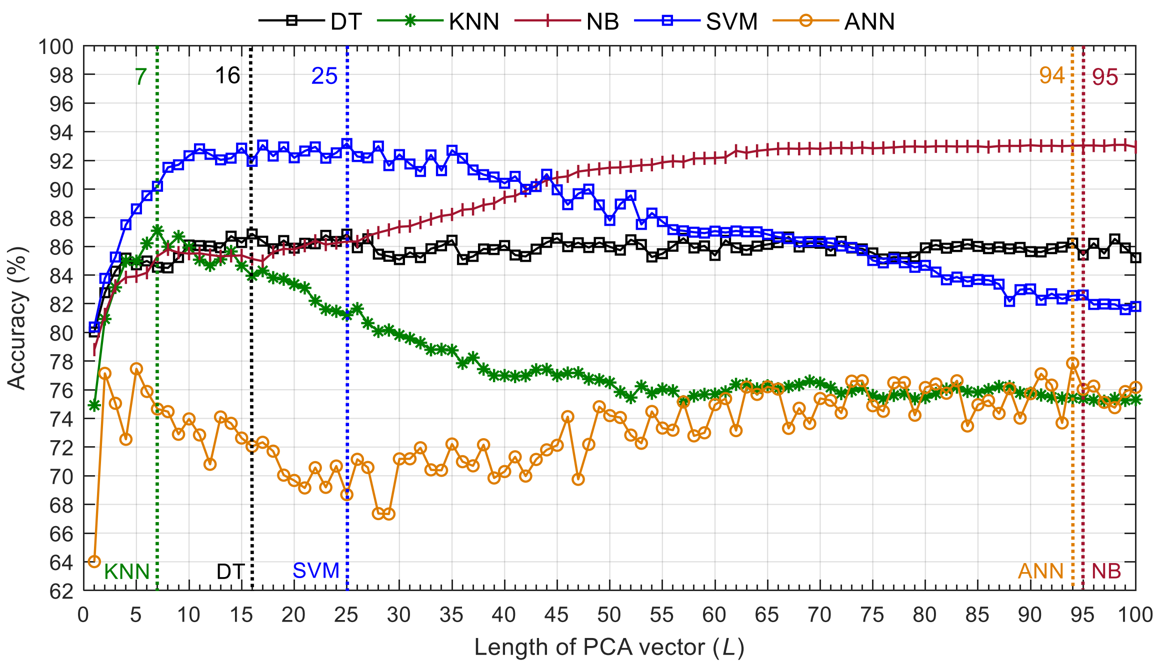

Figure 5. The variation of overall classification accuracy with the length of PCA vectors.

The results in Figure 5 signify that the overall accuracy offered by different classifiers varies significantly in accordance with the length (*L*) of the PCA vectors. It is seen in Figure 5 that the accuracies offered by DT, KNN, NB, SVM and ANN becomes highest for PCA vectors having L = 16, 7, 95, 25 and 94, respectively. Consequently, PCA vectors having these *L* values are finally used to evaluate the performances of corresponding PCA-based classifiers. It is worth to mention that the CNN classifier does not employ PCA and use the raw EEG signals directly for detecting the epileptic seizures.

In this study with 5-fold cross validation, EEG signals involved in each case has been divided in to five folds, i.e., Fold-01, Fold-02, Fold-03, Fold-04 and Fold-05. The classification results of each classifier for each fold have been determined as confusion matrix. For instance, the confusion matrices for each of the five folds for using CNN in 5-class (A-B-C-D-E) classification of EEG signals are shown in Figure 6.

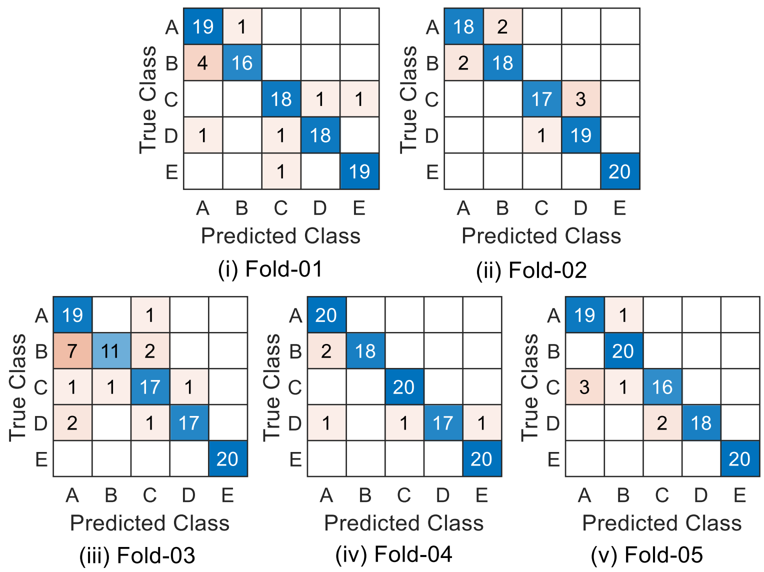

Figure 6. Confusion matrices for using CNN in 5-class classification.

It is observed in each of the confusion matrices in Figure 6 that the proposed CNN classifier can successfully classify EEG signals and can detect epileptic seizures with very few errors. These confusion matrices are then used to compute the classification performances (i.e., accuracy, sensitivity, specificity and precision) of CNN classifier for each class of EEG signals by applying eqns. (7) - (10). For instances, the per-class performances for Fold-01 are presented in Table 3(a). The average performances of CNN classifier for Fold-01 are also included at the bottom of Table 3(a). In a similar fashion, the per-fold performances for each of the five folds are computed which has been shown in Table 3(b). The 5-fold average shown at the bottom of Table 3(b) is ultimately used to evaluate the performances of CNN classifier for 5-class classification in the detection of epileptic seizures.

The results at the bottom of Table 3(b) show that the accuracy, sensitivity, specificity and precision for using CNN classifier in 5-class classification case are 96.32%, 90.80%, 97.70% and 91.60%, respectively. Similarly, we have evaluated the performances of all 2-class, 3-class and 5-class classifiers according to the arrangement of EEG signals listed in Table 1 and Table 2. The classification performances provided by different classifiers for different arrangement of EEG signals are plotted in Figure 7.

Table 3. Classification performances (a) for Fold-01 and (b) per-fold in 5-class classification using CNN classifier.

(a)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
A	94.00	95.00	93.75	79.17
B	95.00	80.00	98.75	94.12
C	96.00	90.00	97.50	90.00
D	97.00	90.00	98.75	94.74
E	98.00	95.00	98.75	95.00
Fold-01 average	96.00	90.00	97.50	90.61

(b)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)
Fold-01-Average	96.00	90.00	97.50	90.61
Fold-02-Average	96.80	92.00	98.00	92.16
Fold-03-Average	93.60	84.00	96.00	86.52
Fold-04-Average	98.00	95.00	98.75	95.49
Fold-05-Average	97.20	93.00	98.25	93.23
All-Fold-Average	**96.32**	**90.80**	**97.70**	**91.60**

The performances of six different classifiers for all classification cases plotted in Figure 7 reveal that the performances of NB, SVM and CNN classifiers are much better than that of other three classifiers (i.e., DT, KNN and ANN). It is observed in Figure 7 that the use of NB, SVM and CNN can provide perfect classification with 100% accuracy, sensitivity, specificity and precision for A-E classification case. For 2-class classification of B-E and C-E cases, the accuracies offered by SVM are 99.50% and 98.50%, respectively. However, the use of both NB and CNN can attain perfect classification performance (i.e., 100%) for these two cases. The classification accuracy achieved by CNN for D-E case is 99.50%, which is significantly better than that achieved by both for using NB (93.00%) and SVM (91.00%). For 2-class classification of ABCD-E case, CNN with 99.80% accuracy also outperforms NB (97.80% accuracy) and SVM (96.80% accuracy). Besides 2-class classification of EEG signals, NB, SVM and CNN also perform better among six different classifiers adopted in this study. The classification accuracies provided NB, SVM and CNN adopted in this study for 2-class classification of EEG signals are comparable to that reported in the research works [12,14,17,23]. For 3-class classification of AB-CD-E case, the accuracies accomplished by NB and SVM are 77.84% and 84.36%, respectively. For such 3-class classification case, CNN can manage to achieve much improved classification accuracy of 98.48%. This classification accuracy is better than 94.46% and 97.07% as reported in the research works [22,23], respectively. For 5-class classification of A-B-C-D-E case, CNN can outperform NB (82.64% accuracy) and SVM (82.08% accuracy) in this study with 96.32% accuracy as shown in Figure 7(a). This 5-class classification accuracy achieved in this study using CNN is also slightly better than 95.84% and 96% accuracies as reported in the research works [23,25]. The performances for all the classification cases presented in Figure 7 confirm that NB, SVM and CNN provides high-performance classification among six different machine learning classifiers adopted in this study. Among these six classifiers, CNN provides maximum classification performances for all cases of 2-class, 3-class and 5-class classification. The highest classification performances achieved in this study are listed in Table 4.

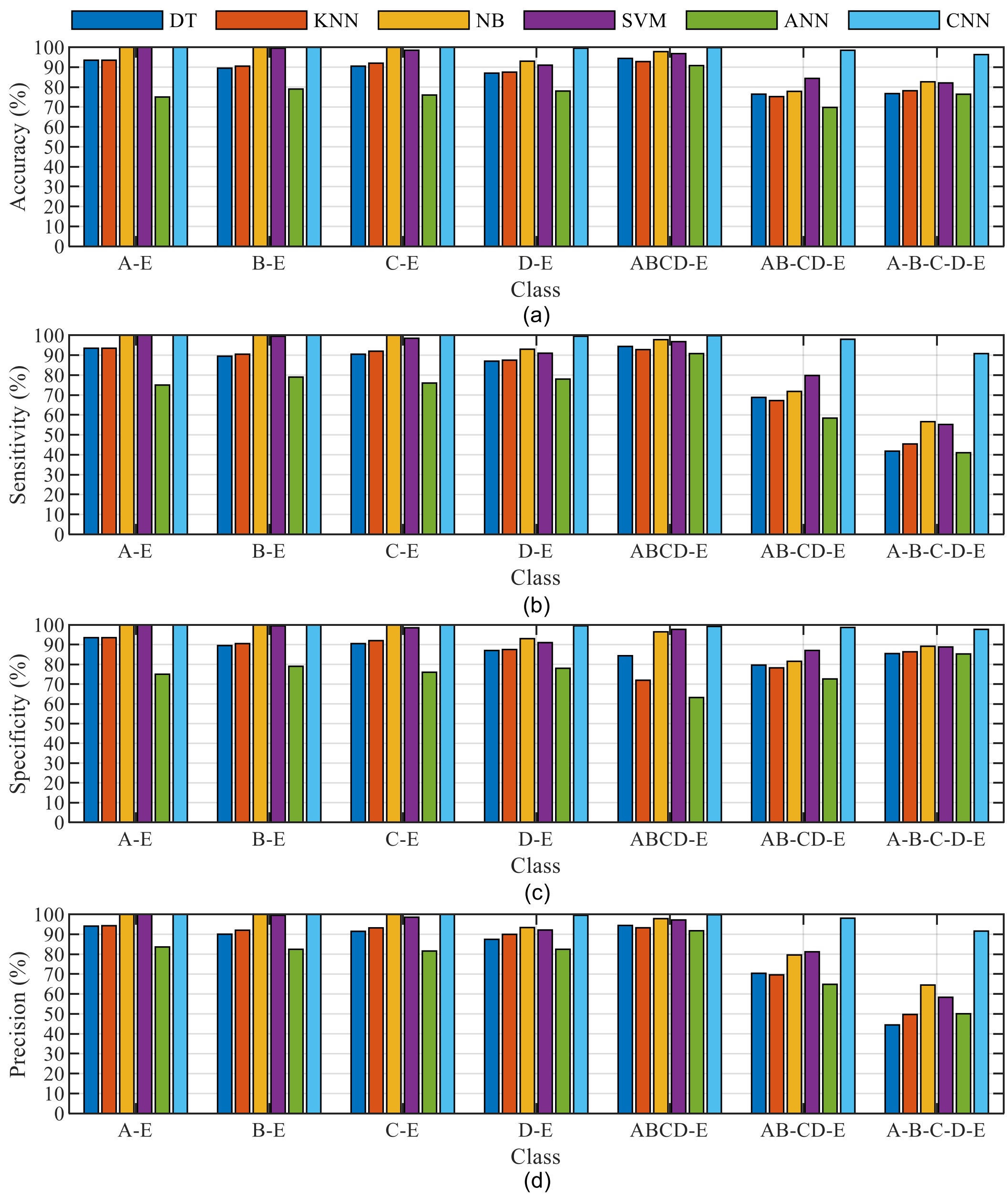

Figure 7. (a) Accuracy, (b) sensitivity, (c) specificity and (d) precision for using different classifiers.

Table 4. Highest classification performances (%) achieved in this study.

Classification	Class	Classifier	Accuracy	Sensitivity	Specificity	Precision
2 - Class	A-E	NB, SVM,	100	100	100	100
	B-E	NB, CNN	100	100	100	100
	C-E	NB, CNN	100	100	100	100
	D-E	CNN	99.50	99.50	99.50	99.52
	ABCD-E	CNN	99.80	99.80	99.20	98.80
3 - Class	AB-CD-E	CNN	98.48	98.00	98.67	98.08
5 - Class	A-B-C-D-	CNN	96.32	90.80	97.70	91.60

The classification performances presented in Table 4 manifest that the customized CNN classifier proposed in this study can attain perfect or near-perfect performances for the classification of EEG signals. This is due to the fact that CNN classifier exploits deep learning algorithm which is enormously capable to extract features from EEG signals. Considering all cases of 2-class, 3-class and 5-class classification of EEG signals, the classification performances achieved in this study have exhibited a reliable and efficient detection of epileptic seizures from EEG signals.

Conclusions

This paper presents a systematic study on the detection of epileptic seizures from EEG signals using machine learning classifiers. The reduced-length PCA vectors have been extracted successfully for the reduction of dimensionality of EEG signals. Then, a total of five machine learning classifiers (i.e., DT, KNN, NB, SVM and ANN) have been used to classify EEG signals by utilizing such PCA vectors. The proper length of PCA vectors to be effectively used by a particular classifier for all 2-class, 3-class and 5-class classification cases is determined systematically. We have also proposed the use of a customized CNN to classify raw EEG signals without applying PCA. The performances of these six classifiers have been analyzed and evaluated in terms of accuracy, sensitivity, specificity and precision. The results show that the use of NB and SVM with PCA can only offer perfect or near-perfect classification performances for some 2-class classification cases, but the classification performances of other PCA-based classifiers are not up to the mark. In addition, the classification performances of all of these PCA-based machine learning classifiers are not satisfactory for 3-class and 5-class classification cases. The results also exhibit that the customized CNN classifier proposed in this study performs best among a total of six different classifiers. The use of such CNN can attain classification performances in the range from 99.50% to 100% for 2-class classification cases. For 3-class case, the proposed CNN can achieve 98.48% accuracy, 98.00% sensitivity, 98.67% specificity and 98.08% precision. Such CNN can also offer much improved classification performances with 96.32% accuracy, 90.80% sensitivity, 97.70% specificity and 91.60% precision for 5-class case. The overall results signify that the proposed CNN classifier can achieve better accuracies irrespective of 2-class, 3-class and 5-class classification cases. Thus, CNN classifier can be an effective tool for the accurate detection of epileptic seizures from EEG signals. We strongly believe that the approach adopted in this study will be a complementary tool to the medical staff for the accurate detection of epileptic seizures from EEG signals.

References

1. England, M.J., Liverman, C.T., Schultz, A.M., &Strawbridge, L. M. (2012). Epilepsy across the spectrum: Promoting health and understanding: A summary of the Institute of Medicine report. Epilepsy & Behavior, 25(2), 266-276.
2. World Health Organization, Epilepsy, Report by the Director-General, 146th session, Provisional agenda item 11, 25 November 2019. https://apps.who.int/gb/ebwha/pdf_files/EB146/B146_12-en.pdf
3. Hussain, M. (2017). Burden of Epilepsy in Bangladesh: Current Approach. Journal of Current and Advance Medical Research, 4(1), 1-2.
4. Mac, T.L., Tran, D. S., Quet, F., Odermatt, P., Preux P.M., &Tan. C.T. (2007). Epidemiology, aetiology, and clinical management of epilepsy in Asia: a systematic review. The Lancet Neurology, 2007, 6(6),533-543.
5. Mohammad, Q.D., Saha, N.C., Alam, M.B., Hoque, S.A., Islam, A., et al. (2020). Prevalence of epilepsy in Bangladesh: Results from a national household survey. Epilepsia Open, 5(4), 526-536.
6. Ahammed, K., & Ahmed, M. U. (2020). Epileptic Seizure Detection Based on Complexity Feature of EEG. Journal of Biomedical Analytics, 3(1), 1–11.
7. Wang, L., Xue, W., Li, Y., Luo, M., Huang, J., Cui, W., & Huang, C. (2017). Automatic epileptic seizure detection in EEG signals using multi-domain feature extraction and nonlinear analysis. Entropy, 19(6), 222.
8. Yan, T., Dong, X., Mu, N., Liu, T., Chen, D., Deng, L., Wang, C., & Zhao, L. (2018). Positive classification advantage: tracing the time course based on brain oscillation. Frontiers in human neuroscience, 11, 659.
9. Nahzat, S., & Yaganoglu, M. (2021). Classification of Epileptic Seizure Dataset Using Different Machine Learning Algorithms and PCA Feature Reduction Technique. Journal of Investigations on Engineering & Technology, 4(2), 47-60.
10. Sahu, R., Dash, S. R., Cacha, L, A., Poznanski, R. R., & Parida. S. (2020). Epileptic seizure detection: a comparative study between deep and traditional machine learning techniques. Journal of Integrative Neuroscience, 19(1), 1-9.
11. Almustafa, K. M. (2020). Classification of epileptic seizure dataset using different machine learning algorithms. Informatics in Medicine Unlocked, 21, 100444.
12. Jaglan, S., Dhull, S. K., &Singh, K. K. (2021). Tertiary wavelet model based automatic epilepsy classification system. International Journal of Intelligent Unmanned Systems, ISSN: 2049-6427.
13. Aayesha, Qureshi, M.B., Afzaal, M., Qureshi, M.S., & Fayaz, M. (2021). Machine learning-based EEG signals classification model for epileptic seizure detection. Multimedia Tools and Applications 80, 17849–17877.
14. Sharmila, A., & Geethanjali, P. (2016). DWT Based Detection of Epileptic Seizure from EEG Signals Using Naive Bayes and k-NN Classifiers. IEEE Access, 4, 7716-7727.

15. Savadkoohi, M., Oladunni, T.,& Thompson, L. (2020). A machine learning approach to epileptic seizure prediction using Electroencephalogram (EEG) Signal. Biocybernetics and Biomedical Engineering, 40(3), 1328-1341.
16. Wang, Z., Na, J., & Zheng, B. (2020). An Improved kNN Classifier for Epilepsy Diagnosis. IEEE Access, 8, 100022-100030.
17. Jaiswal, A. K., & Banka, H. (2018). Epileptic seizure detection in EEG signal using machine learning techniques. Australas Physical & Engineering Sciences in Medicine, 41(1), 81-94.
18. Tawfik, N. S., Youssef, S. M., & Kholief, M. (2016). A hybrid automated detection of epileptic seizures in EEG records. Computers & Electrical Engineering, 53, 177-190.
19. Shankar, R. S., Raminaidu, C. H., Raju, V. V. S., & Rajanikanth, J. (2021). Detection of Epilepsy based on EEG Signals using PCA with ANN Model. Journal of Physics: Conference Series, 2070, 012145.
20. Baykara, M., & Abdulrahman, A. (2021). Seizure detection based on adaptive feature extraction by applying extreme learning machines. Traitement du Signal, 38(2), 331-340.
21. Srinivasan, V., Eswaran, C., & Sriraam, N. (2005). Artificial Neural Network Based Epileptic Detection Using Time-Domain and Frequency-Domain Features. Journal of Medical Systems, 29(6), 647-660.
22. Liu, Y. H., Chen, L., Li, X. W., Wu, Y. C., Liu, S., Wang, J. J., Hu, S. G., Yu, Q., Chen, T. P., & Liu, Y. (2022). Epilepsy detection with artificial neural network based on as-fabricated neuromorphic chip platform. AIP Advances 12, 035106.
23. Zhao, W., & Wang, W. (2020). SeizureNet: a model for robust detection of epileptic seizures based on convolutional neural network. Cognitive Computation and Systems, 2(3), 119-124.
24. Chirasani, S.K.R., &Manikandan, S. (2022). A deep neural network for the classification of epileptic seizures using hierarchical attention mechanism. Soft Computing, 26, 5389–5397.
25. Tuncer, T., Dogan, S., Naik, G.R., & Pławiak, P. (2021). Epilepsy attacks recognition based on 1D octal pattern, wavelet transform and EEG signals. Multimed Tools and Applications, 80, 25197–25218.
26. Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P. & Elger, C.E. (2002). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Physical review. E. 64, 061907.
27. https://www.ukbonn.de/en/epileptology/workgroups/lehnertz-workgroup-neurophysics/downloads(accessed on 12 July 2022)
28. Azad, A.K., Khan, F. N., Alarashi, W. H., Guo, N., Lau, A. P. T., & Lu, C. (2017). Temperature extraction in Brillouin optical time-domain analysis sensors using principal component analysis based pattern recognition. Optics Express, 25(14),16534-16549.