CLASSIFICATION OF ICTAL AND INTERICTAL EEG SIGNALS

Zavid M. Parvez and Manoranjan Paul
School of Computing and Mathematics
Charles Sturt University
Bathurst, Australia
Email: mparvez@csu.edu.au and mpaual@csu.edu.au

ABSTRACT
An electroencephalogram (EEG) is a graphical record of ongoing electrical activity produced by firing of neurons of the human brain due to internal and/or external stimuli. Feature extraction and classification of the EEG signals are used for diagnosis the epileptic seizure (i.e., physical changes in behaviour that occur due to abnormal electrical activity in the brain). Classification of Ictal (i.e., seizure period) and Interictal (i.e., interval between seizures) EEG signals is very important for the treatment and precaution of an epileptic patient. However, the classification accuracy of Ictal and Interictal EEG signals is not at satisfactory level due to their non-abruptness phenomenon using the existing seizure and non-seizure classification methods. Moreover, the features of Ictal and Interictal signals are not consistence in different locations for an epileptic period. In this paper we present new approaches for features extraction of Ictal and Interictal using various transformations such as discrete cosine transformation (DCT), DCT-discrete wavelet transformation, and singular value decomposition. The least square support vector machine is applied on the features for classifications. Results demonstrate that our proposed methods outperform the existing state-of-the-art method in terms of classification accuracy for the large benchmark dataset in different brain locations.

KEY WORDS
EEG signals, Seizure, Epilepsy, Ictal, Interictal, EMD, SVD, and LS-SVM.

1. Introduction

The human brain is one of the complex organs in the entire body. Brain consists with approximately 100 billion, perhaps even more, nerve cells called neurons. Interconnected neurons of the brain are responsible for higher order thinking and complex behaviour. Human brain is an organic electrochemical computer as neurons exploit chemical reaction to generate electricity. Electrochemical nature of neurons gives to rise up our actions, our modes and our behaviour. When neuron is excited by external/internal stimuli, it transfer electrochemical impulses incoming from the dendrites along the axon and onto the next neuron [1] shown in Fig 1. Electroencephalogram (EEG) measures the changes of the electrical activity in term of voltage fluctuations of the brain within short period of time, usually 20-40 minutes, through multiple electrodes place on the scalp [2]. In the clinical contexts, the main diagnosis of EEG is to discover abnormalities of brain activity refer to the epileptic seizure. A seizure occurs when the neurons generate uncoordinated electrical discharges that spread throughout the brain and epilepsy is a recurrent seizure disorder caused by abnormal electrical discharges from brain cells, often in the cerebral cortex [3]. Another clinical use of EEG is in diagnosis of coma, brain death, and encephalopathies. Moreover, EEG can be used in many applications such as emotion recognition [4], video quality assessment [5], alcoholic consumption measurement [6], sleep stage detection [7], change the brainwaves by smoking [8] and mobile phone usages [9], etc.

Fig 1: Structure of the neuron [1].

People may experience abnormal activities in sensation, movement, awareness, and behaviour during seizure as a result they cannot perform normal task. Ictal and Interictal both are medical condition of seizure where Ictal represents the period of seizure and Interictal represents the intermediate period between two seizures. Note that Interictal is significantly different from normal non-seizure signal. A prediction of seizures (i.e., Ictal) from Interictal could aware a patient to put away from the next seizure. A number of existing EEG classification techniques [11]-[17] is able to classify seizure and non-seizure EEG signals for a small size dataset almost perfectly. However, those methods are struggling to
provide acceptable classification accuracy for Ictal and Interictal EEG signals due to the non-abruptness phenomenon [10] of the signals.

Existing feature extraction and classification methods based on wavelet [11]- [13] and Fourier transformation [14] were employed for the detection of seizure in EEG signals. Panda et al. [11] computed various features like energy, entropy, and standard deviation by discrete wavelet transform and used support vector machine (SVM) as a classifier. Dastidar et al. [12] applied wavelet transformation to decompose the EEG signals into different range of frequencies and three features, such as standard deviation, correlation dimension, and the largest Lyapunov exponent (quantifying the non-linear chaotic dynamics of the signals), are employed and different methods applied for classification. Ocak [13] proposed fourth level wavelet packet decomposition method to decompose the normal and epileptic EEG epochs to various frequency bands and then used genetic algorithm to find optimal feature subsets which maximize the classification performance. Polat et al. [14] proposed two stage processes: first was feature extraction using first Fourier transform (FFT) and second was decision making using decision making classifier. The techniques [11] [12] [14] have used small dataset [18] (for detailed information of the dataset, see Section 2.1). Sample examples of the dataset are shown in Fig 2. The classification accuracies were 91.2% [11], 96.7% [12], and 98.72% [14], respectively. Ocak’s [13] experimental dataset was the same pattern with the dataset in [18] and the classification accuracy was 98%.

Fig 2: Seizure and non-seizure dataset [18], first four rows and fifth row indicated that the non-seizure and seizure data, respectively.

Other feature extraction and classification methods based on empirical mode decomposition (EMD), principle component analysis (PCA), and genetic algorithm (GA) have been proposed which are mentioned in [15]-[17]. Liang et al. [15] extracted EEG features where the approximate entropy (ApEn) analysis was evaluated for its ability to analyse multiclass EEGs, and the performance of ApEn analysis was enhanced by incorporating the power of EEG sub-bands or autoregressive models. PCA and GA were compared for their ability to select features by applying various linear and nonlinear methods for classification, accuracy was maximum 98.67%. Pachori [16] decomposed EEG signals into intrinsic mode function (IMF) using empirical mode decomposition (EMD) and then computed mean frequency for each IMF by Bessel function [21] to differentiate seizure and non-seizure EEG signals. Zhang et al. [20] proposed an approach for the pattern recognition of four complicated hand activities, such as grasping a football, a small bar, a cylinder and a hard paper, based on EEG signal, in which each piece of raw data sequence for EEG signal is decomposed by wavelet transform to form a matrix and then singular value decomposition (SVD) is applied on the matrix to find the singular value feature vectors. After that these features are used as an input to artificial neural network to discriminate four hand activities and got 89% classification accuracy. Bajaj et al. [17] extracted amplitude modulation and frequency modulation as features by EMD. Among the existing contemporary techniques, Bajaj et al.’s technique is the latest and the best in terms of performance. They used least squares-SVM (LS-SVM) [22] technique for the classification of seizure and non-seizure EEG signals using the small dataset [18]. They obtained 98.5% accuracy using radial basis function (RBF) kernel in LS-SVM classifier.

Existing methods [11]-[17] except [13] used small dataset [18] (for detailed information of the dataset, see Section 2.1). The small dataset in [18] with duration 23.6 second has seizure (i.e., Ictal) and non-seizure signals which can be distinguished by their visual phenomena such as magnitude of amplitude and changing rate of frequency (see in Fig 2). For the non-seizure signal the amplitude is low and the frequency is high while the nature of seizure signal is totally reverse (see Fig 2). In this experiment, we have used a large dataset [19] for comparing Ictal and Interictal EEG signals with duration 660 seconds. We have used two different locations: Frontal and Temporal lobe of the human brain. See sample signals in Fig 3 and

![Fig 3: Ictal (Patient one, 7th datablock, channel one) and Interictal (Patient one, 95th datablock, channel one) of Frontal Lobe dataset from Epilepsy Center of the University Hospital of Freiburg [19].](image-url)
The detailed explanation of the dataset is given in Section 2.1. Fig 3 and Fig 4 demonstrated the non-abrupt phenomena (i.e., not easily distinguishable between Ictal and Interictal based on amplitude and frequency) of the Ictal and Interictal signal for both cases of Frontal and Temporal lobe compared to the dataset in [18]. Moreover, EEG signals from different locations exhibit different phenomenal activities for an Ictal and Interictal period. Thus, classification of Ictal and Interictal EEG signals are challenging compared to that of seizure (i.e., Ictal) and non-seizure EEG signals.

The data obtained by Neurofile NT digital video EEG system with 128 channels, 256 Hz sampling rate, and 12-bit A/D resolution. The recorded EEG data were acquired through surface electrodes placed on the skull using the international 10–20 electrode placement scheme i.e., 32 electrodes [14]. Among them two subsets are taken from health volunteers, two subsets are seizure free intervals and another subset has also taken during seizure period from five patients [18] (see sample examples in Fig 2).

The paper is organized as follows: the detailed of the proposed three methods with dataset, feature extractions, and classifiers are described in section 2; the elaborate experimental results are explained in section 3 while section 4 concludes the paper.

2. Proposed Methods

We propose three feature extraction methods based on well established transformations/decompositions such as DCT, DWT, and SVD. In the first proposed method, we extracted entropy and energy from the high frequency DCT coefficients because we have observed that the energy and entropy of the high frequency DCT coefficients of Ictal and Interictal signals are significantly different. In the second proposed method, we extract features from the standard deviation of high frequency DCT coefficients and high frequency of DWT coefficients. In the third proposed method, we extract features from the standard deviation of the higher and lower portions of the singular values by applying SVD. We use LS-SVM classifier and the large dataset [19] with Ictal and Interictal signal of 12 patients captured from Frontal and Temporal lobe for results comparison against the state-of-the-art method [17].

2.1 Dataset

The existing seizure and non-seizure feature extraction and classification schemes [11]-[17] except [13] used small dataset [18]. The dataset consists of five subsets, each subset contains 100 single-channels recording, and each recording has 23.6 s in duration captured by the international 10–20 electrode placement scheme i.e., 32 electrodes [14]. Among them, two subsets are taken from healthy volunteers, two subsets are seizure free intervals and another subset has also taken during seizure period from five patients [18] (see sample examples in Fig 2).

Ocak [13] used different dataset for his experiment obtained from five normal subjects and five epileptic patients. The EEG data were acquired through surface electrodes placed on the skull using the international 10–20 electrode placement scheme [14]. There were a total of 400 single channel EEG epochs where 300 were normal and 100 were epileptic. Each epoch consisted of a 23.6 s long EEG recording. The recordings were collected using a 128-channel amplifier system and digitized at 173.61 Hz sampling rate and 12-bit A/D resolution.

The dataset used in the paper was recorded at Epilepsy centre of the University Hospital of Freiburg, Germany [19]. The data obtained by Neurofile NT digital video EEG system with 128 channels, 256 Hz sampling rate,
and 16 bit analogue-to-digital converter. Data recording temporarily paused after each block due to technical reasons and pause time 1-3 seconds. For each of the patients, there are datasets called Ictal and Interictal. Firstly, system contains epileptic seizures file with 50 min pre-ictal data. Later on, it was containing 24h of EEG recording without seizure activity. The ictal periods were determined based on identification of typical seizure patterns of experienced epileptologists. In the experiment, we used Ictal and Interictal dataset of Frontal and Temporal Lobe of twelve patients with 660 seconds duration which is 28 times longer than the small dataset [18]. The total sample length of each EEG signal is around 900,000 with duration of 24 hours. For the experiment we have taken 6,400 samples from 10,001 to 16,400 of each EEG signal with 660 seconds duration.

We have cut off the noisy data which have excessive amplitude. A sample is considered as an outlier if the absolute value of the sample is higher than the three times of standard deviation of the data. The outlier noisy data is modified as the three times of standard deviation of the data. Two sample signals of the dataset are shown in Fig 3 and Fig 4 after modifying the signals. The existing of the noisy data in the dataset [19] is mentioned in the metadata.

2.2 Feature Extraction using DCT

In our experiment, we apply DCT on benchmark dataset [19] and take last 25% of DCT coefficient (i.e., higher frequency components). These frequency components generate statistical feature such as entropy and energy which can distinguish Ictal and Interictal signals. We defined entropy and energy of the high frequency coefficients in the similar fashion used in [11] as:

\[
\text{Entropy} = - \sum_{i=1}^{n} X_i(t) \ln(X_i^2(t)) \quad (1)
\]

\[
\text{Energy} = \sum_{i=1}^{n} |X_i(t)|^2 \quad (2)
\]

where \(n\) is the length of signal, \(X(t)\) is the higher frequency components of the DCT coefficients applied on the original signal \(x(t)\). The representation of Ictal and Interictal data for different higher frequency DCT coefficients (i.e., 25% of total DCT coefficients) and their entropy and energy are given in Fig 5. It can be observed that the changing rate of frequencies is higher for Interictal signal than Ictal signal from Temporal. It is also interesting to observe that the amplitude of the higher frequency DCT coefficients for Ictal signals is larger compared to that of Interictal signals. Fig 5 also demonstrates that the entropy and energy the extracted from the higher frequency DCT components have different values for Ictal and Interictal signals. Thus, entropy and energy would be effective features for Ictal and Interictal signals classification.

2.3 Feature Extraction using DCT-DWT

First we apply DCT and DWT on the EEG signals separately. Then, we calculate the standard deviation from the last 25% of DCT coefficients (i.e., high frequency) and all high frequency components of DWT. The equation of standard deviation calculation is as follows:

\[
\sigma_{std} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{x})^2} \quad (3)
\]

where \(X_i\) is the DCT or DWT coefficients of the sample items and \(\bar{x}\) is the mean value of the coefficients \(X_i\).

Two standard deviation features from DCT and DWT, higher frequency components are then used for classification. The wavelet transform can better localize the signal components in time frequency space. The Morlet-Grossmann defined the continuous wavelet transform for one dimensional \(x(t)\) signal as follows [23]:

\[
W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi^*(\frac{t-b}{a})dt \quad (4)
\]

where \((\cdot)^*\) denoted the complex conjugate, \(\psi(t)\) is the analysing wavelet, \(a\) is the scalar parameter, and \(b\) is the position parameter.

The most common method is to separate continuous wavelet to discrete wavelet by scale factor of binary scales. Let \(a = 2^j\) and the equation is as follows [20]:

\[
W(j,b) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} x(t) \psi^*(\frac{t-b}{2^j})dt \quad (5)
\]

The first split of signal \(x(t)\) is \(a_1\) and \(d_1\). Then further \(a_1\) is split into \(a_2\) and \(d_2\). This procedure continues until
specified level is reached. For example, \( a_i \) and \( d_i \) \((i = 1, 2, \ldots, n)\) are the low frequency component and high frequency component of the \( i^{th} \) level respectively. In the experiment we have taken only the high frequency DWT components. The first two rows in Fig 6 show the higher change of frequency for Interictal compared to Ictal from Temporal lobe. The third row of Fig 6 represents that the standard deviation generated from the high frequency component of DCT and DWT for Ictal is higher than that of Interictal signal. Thus, the standard deviations extracted from the DCT and DWT coefficients can be good features for classifications.

\[
A = U \Sigma V^H
\]

where \( A_{mm} \) is a matrix, \( U_{mm} \) and \( V_{mm} \) are orthogonal matrix, and \( \Sigma_{mm} \) is a matrix with non-negative diagonal entries. The columns of \( U \) are call left singular vectors and the rows of \( V^H \) are called right singular vectors. If \( A \) is a rectangular \( m \times n \) matrix of rank \( K \) than \( U \) will be \( m \times m \) and \( \Sigma \) will be:

\[
\begin{bmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \lambda_2 & \cdots & 0 \\
0 & 0 & \ddots & \cdots \\
0 & 0 & \cdots & \lambda_n
\end{bmatrix}
\]

where \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{m,1} \geq \lambda_{n,n} \).

The data distribution of non-zero singular values of the Interictal signal is bigger than that of the Ictal signal from the dataset of Frontal lobe shown in the top part of Fig 7. The bottom part of Fig 7 shows the standard deviation extracted from the 25% of non-zero singular values from Ictal and Interictal signals respectively. The figure reveals that the standard deviations extracted from the Ictal and Interictal signals have distinguishable values, thus, they can be used for Ictal and Interictal signal classification.

![Fig 6: High frequency DWT coefficients and standard deviation of high frequency DCT and DWT coefficients for the EEG signals of Temporal lobe; the first two rows show that the high frequency DWT coefficients of Ictal and Interictal signals, while the last row shows the standard deviation of the high frequency DCT and DWT coefficients of Ictal and Interictal.](image)

![Fig 7: Non-Zero singular value of SVD for Frontal lobe. The first row shows that the non-zero singular value distribution for Ictal and Interictal signals from Frontal lobe and the last row shows the standard deviation of the first and the last 25% non-zero singular values of Ictal and Interictal signals.](image)

### 2.4 Feature Extraction using SVD

To extract the features using SVD, we have to reshape the EEG signals into a two dimensional matrix. The column vector of EEG signal is rearranged to square matrix for the computation of singular value components using SVD. From the singular values, we take the first and the last 25% of the non-zero diagonal values to calculate the standard deviation and treated as the two features to classify Ictal and Interictal signals. The SVD allows transforming any given matrix \( A \in C^{mn} \) to diagonal form using unitary matrices, i.e.

\[
A = U \Sigma V^H
\]

To classify the Ictal and Interictal signals, we have extracted various features from DCT, DWT, and SVM. For classification, we have used SVM-based classifier as the SVM is one of the best classifiers in the EEG signal analysis. SVM is a potential methodology for solving problem in linear and nonlinear classification, function estimation, and kernel based learning methods [24]. It can minimize the operational error and maximize the margin hyperplane, as a result it will maximize the classification performance [24]. LS-SVM is the extended version of SVM and it is closely related to regularization networks and Gaussian process and it has primal-dual interpretations [22]. The equation of LS-SVM can be defined in [17] as:

\[
f(x) = \text{sign} \left[ \sum_{i=1}^{N} \alpha_i y_i K(x, x_i) + c \right]
\]

where \( K(x, x_i) \) is a kernel function, \( \alpha_i \) are the Lagrange multipliers, \( c \) is the bias term, \( x_i \) is the training input, and \( y_i \) is the training output pairs. RBF and polynomial kernel are used in our experiments and these functions can be defined in [17] as:

\[
k(x, x_i) = \exp \left( -\frac{\|x-x_i\|^2}{2\sigma^2} \right)
\]

where \( \sigma \) controls the width of RBF kernel function.
\[ k(x, x_i) = (x^T x_i + 1)^L \]  \hspace{1cm} (10)

where \( L \) is the degree of polynomial.

3. Experimental Results

Our ultimate target is to classify Ictal and Intercital EEG signals using LS-SVM with RBF and polynomial kernel where the value of regularization parameter and kernel parameter are 10 and 6, respectively. In this paper, three methods have been proposed based on the new features of Ictal and Intercital EEG signals captured from Frontal and Temporal lobe using DCT, DWT and SVD. In our experiment, we have used 200 signals for Ictal and 800 signals for Intercital, each signal has 660 seconds recording data. After feature extraction, we randomly sorted the Ictal and Intercital data for the 1000 trials for each lobe either Frontal or Temporal. For unbiased classification results [25], we divide the whole 1000 trials into four subsets, in each sub-set we randomly select 75% samples for the training and the remaining 25% for the testing. After that, we calculate average value using the four sub-sets to compute sensitivity, specificity and accuracy [17]. The sensitivity, specificity and accuracy are defined as:

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \times 100 \hspace{1cm} (11)
\]

\[
\text{Specificity} = \frac{TN}{TN + FP} \times 100 \hspace{1cm} (12)
\]

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \hspace{1cm} (13)
\]

where TP and TN represents the total number of detected true positive events and true negative events respectively. The FP and FN represent false positive and false negative respectively.

Table 1 shows the sensitivity, specificity and accuracy which is referred to classification performance for different features of Ictal and Intercital EEG signals from Frontal and Temporal lobe using LS-SVM classifier with RBF and Polynomial kernel function. It can be observed for RBF kernel from Table 1 that the features extracted from SVD has the best classification accuracy i.e., 89.3% for the dataset from Frontal lobe and feature extracted from DCT-DWT has the best accuracy i.e., 85.8% for the dataset of Temporal lobe. Moreover, for polynomial kernel from Table 1, it can be observed that SVD has better classification accuracy i.e., 88.8% for the dataset from Frontal lobe and DCT-DWT has 84.8% classification accuracy for Temporal lobe dataset. It can be concluded for both kernel of LS-SVM classifier that the features extracted from DCT-DWT or SVD outperform the state-of-the-art method [17] in terms of above mentioned three performance criteria.

![Receiver Operating Characteristic curve](image)

Fig 8: The receiver operating characteristics (ROC) curve of third subset of training and testing EEG signals using LS-SVM with RBF kernel from Frontal lobe.

The performance of the LS-SVM is evaluated by receiver operating characteristics (ROC) plots is shown in Fig 8. ROC illustrates the performance of a binary classifier system where it is created by plotting the fraction of true positives from the positives i.e., true positive rate (TPR) vs. the fraction of false positives from negatives i.e., false positive rate (FPR) with various threshold settings. TPR is known as sensitivity, and FPR is one minus the specificity or true negative rate. The Fig 8 demonstrated that DCT-DWT and SVD are carrying good classification results than that of start-of-the-art method [17].

Fig 9 shows the distribution of normalized feature values (i.e., standard deviation data using DCT, DWT, and SVD) of Ictal and Intercital from the Frontal lobe dataset [19]. The figure shows that the data distribution for Ictal is higher than that of Intercital signals which is the other
evidence why we use those features to classify Ictal and Interictal EEG signals in the proposed methods.

Fig 9: Normalized distribution of feature values extracted from Ictal and Interictal EEG signals using DCT, DWT, and SVD where 200 samples are taken from each of Ictal and Interictal categories in Frontal lobe.

Fig 10: Classification of Ictal and Interictal signals from Frontal lobe for the third subset of training and testing using the method [17] based on EMD.

Fig 10, Fig 11, Fig 12, and Fig 13 show the classification comparison using the proposed methods and the state-of-the-art method [17] visually for better understanding. Fig 10 and Fig 11 show LS-SVM classification results using EMD and SVD respectively. We consider Fig 10 and Fig 11 from Frontal lobe for the third subset of training and testing and obtain classification accuracy 84.8% for SVD and 77.2% for EMD. It can be concluded from the figures that the proposed method based on the SVD outperforms the state-of-the-art method based on the EMD.

The same classification comparisons are also provided in Fig 12 and Fig 13 using the state-of-the-art method based on the EMD and the proposed method based on the DCT-DWT respectively. We consider Temporal lobe for the fourth subset of training and testing, obtained classification accuracy 96.4% for DCT-DWT and 85.2% for EMD (see Fig 12 and Fig 13). Thus, it is concluded that the proposed method based on the DCT-DWT also outperforms the EMD method.

Fig 12: Classification of Ictal and Interictal signals from Temporal lobe for fourth subset of training and testing used EMD method.

Fig 13: Classification of Ictal and Interictal signals from Temporal lobe for fourth subset of training and testing used DCT-DWT method.
4. Conclusion

Unlike seizure and non-seizure EEG signals, there is non-abrupt phenomena between Ictal and Intercital EEG signals. Moreover, the features of Ictal and Intercital signals are not consistence in different locations for an epileptic period of a patient. Thus, the classification performance for Ictal and Intercital EEG signals using the existing features extraction techniques is not at expected level. To improve the accuracy, we proposed three new feature extractions methods in this paper based on DCT, DCT-DWT, and SVD to classify Ictal and Intercital EEG signals. From DCT, entropy and energy features have been extracted from high frequency DCT coefficients. From DCT-DWT, standard deviation features have been extracted from the high frequency components of DCT and DWT. From SVD, standard deviation features have been extracted from the larger and smaller singular values. The experimental results show that the proposed methods with the extracted features from DCT, DCT-DWT, and SVD with LS-SVM classifier outperform the state-of-the-art method based on EMD-based features in terms of accuracy, sensitivity, and specificity for the Ictal and Intercital EEG signals from the large benchmark dataset in two different brain locations.

References


[19] EEG Data set from Epilepsy Center of the University Hospital of Freiburg http://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/egge- database. Visited Date: June 10, 2012.


