DWT Based Transformed Domain Feature Extraction Approach for Epileptic Seizure Detection

by

Mahajabin Mostafa
18101458
Mohtasim Abrar Samin
18101094
Nabila Bintey Hassan
18101237
Saiara Zerin Nibras
18101251
Samir Rahman
18101130

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
September 2021

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Mahajabin Mostafa
18101458

Mohtasim Abrar Samin
18101094

Nabila Bintey Hassan
18101237

Saiara Zerin Nibras
18101251

Samir Rahman
18101130

Approval

The thesis/project titled "DWT Based Transformed Domain Feature Extraction Approach for Epileptic Seizure Detection" submitted by

1. Mahajabin Mostafa (18101458)
2. Mohtasim Abrar Samin (18101094)
3. Nabila Bintey Hassan (18101237)
4. Saiara Zerin Nibras (18101251)
5. Samir Rahman (18101130)

Of Summer, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on September 26, 2021.

Examining Committee:

Supervisor:
(Member)

Zavid Parvez
Digitally signed by Zavid Parvez
DN: cn=Zavid Parvez, o=Brac University, ou=CSE, email=zavid.parvez@bracu.ac.bd, c=US
Date: 2021.10.05 16:53:36 +11'00'

Mohammad Zavid Parvez, PhD
Assistant Professor
Department of Computer Science and Engineering
BRAC University

Secondary Supervisor:
(Member)

Mohammed Abid Abrar
Lecturer
Department of Computer Science and Engineering
BRAC University

Program Coordinator:
(Member)

Md. Golam Rabiul Alam, PhD
Assistant Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

Hereby, we, the members, consciously assure that for the manuscript /insert title/ the following is fulfilled:

- This content is the writers' own unique work, which has never been published before.
- The work accurately and completely represents the authors' own research and analysis.
- The study appropriately acknowledges co-authors and co researchers for their significant contributions.
- The findings are suitably contextualized in light of previous and ongoing research.
- All sources are appropriately mentioned (correct citation). Text that is literally copied must be identified as such by using quote marks and providing suitable reference.
- All authors were personally and actively involved in significant effort leading up to the article and will accept public responsibility for its content.

Violations of the Ethical Statement standards may have serious repercussions. We agree with the preceding declarations and certify that this submission adheres to BRAC University's rules.

Abstract

Epileptic seizure is a neurological disorder that is prevalent in both males and females of all age ranges. Detection of epileptic seizure serves as an important role for epileptic patients as it allows the initiation of systems to prevent injuries and limiting the possibilities of risk by providing targeted therapy by anticipating their onset prior to presentation. Electroencephalogram (EEG) plays an important role in seizure detection and is one of the most well-known techniques for determining stages of epilepsy. Since, EEG is a non-stationary signal it can be quite difficult to differentiate amongst seizure activity and normal neural activity. In this paper we have proposed an epilepsy detection method based on five different feature extraction methods and followed by that the original domain of the extracted features were transformed using Discrete Wavelet Transform (DWT) and three different classifiers- Decision Tree, Random Forest and KNN to classify into seizure and non-seizure stages. Results demonstrated in this paper have outperformed the existing state-of-the-art methods with 97.22%, 100% and 83.33% for 2 class classification and 91.67%, 91.67% and 80.56% for 4 class classification for the aforementioned classification techniques accordingly.

Keywords: DWT, Transformed Domain, EEG, Feature Extraction, Classification

Dedication

This thesis is dedicated to our institution's mentors, without whom we would not have been able to complete this thesis. They not only provided us with academic knowledge, but they also provided us with vital guidance when we needed it the most.

Acknowledgement

First and foremost, we express our gratitude to Almighty Allah for allowing us to complete our Thesis on time and without any obstacles.
With that being said, we'd like to express our sincere appreciation to Dr. Mohammad Zavid Parvez, our distinguished instructor and supervisor, for his steadfast support and persistent supervision, which enabled us to finish our project. We owe a debt of appreciation to Mohammed Abid Abrar, our secondary supervisor, for generously assisting us with this thesis to the best of his abilities.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

ANN Artificial Neural Network

$ApEn$ Approximate Entropy

csv Comma Separated Values

DCT Discrete Cosine Transform

DE Embedding Dimension

DFA Detrended Fluctation Analysis

DTC Decision Tree Classifier

DWT Discrete Wavelet Transform

ECG Electrocardiogram

EEG Electroencephalogram

EMD Empirical Mode Decomposition

FB Fourier-Bassel

FD Fractal Dimension

FFT Fast Fourier Transform

GA Genetic Algorithm

HFD Higuchi Fractal Dimension

IMF Intrinsic Mode Function

$KLDivergence$ Kullback–Leibler Divergence

KNN K-Nearest Neighbor

$LS-SVM$ Least Square- Support Vector Machine

LVQ Learning Vector Quantization

MF Mean Frequency

$P - value$ Probability Value

PCA Principle Component Analysis

RBF Radial Basis Function

RFC Random Forest Classifier

RMS Root Mean Square

SVD Singular Value Decomposition

SVM Support Vector Machine

$t - SNE$ t-distributed Stochastic Neighbor Embedding

VNS Vagal Nerve Stimulus

WHO World Health Organization

WT Wavelet Transform

Chapter 1

Introduction

1.1 Motivation

Epilepsy is defined as a neurological disorder further characterized by the electrical disturbance and uncoordinated electrical discharges in our brain leading to repeated malfunction or recurrent seizures. The term seizure specifies this disorder causing sudden movements, convulsions and the loss of consciousness. According to the WHO, around 50 million people have been affected with epilepsy worldwide with both adults and children falling under the diagnosis range. In addition to that, it is one of the oldest recognized conditions in the history of neurology, making it a globally prevalent neurological disease[1]. The Epilepsy Therapy Project revealed that around 10 percent of people around the world will experience seizure at some point of their life, noting down a staggering statistic which said that 1 in 26 people may develop this neurological disorder from the seizure experienced[2]. Patients experiencing epileptic seizures may be subjected to uncountable injuries such as fractures, submersion, burns, motor vehicle accidents and the worst possible case, death. Therefore, detecting seizures beforehand prevents countless injuries while rapidly promoting the therapies aimed at treating seizures and increasing the importance of seizure detection greatly. So, in this paper we have proposed an epilepsy detection method based on five different feature extraction methods. Following that, the original domain of the extracted features was transformed using DWT and three different classifiers- Decision Tree, Random Forest and KNN to classify into seizure and non-seizure stages.

1.2 History of Epilepsy

The history of epilepsy and its treatment traces back to ancient civilization, during 2000 B.C.[3] The diagnosis and treatment of this condition during the past centuries have been medically inaccurate to great extent, either being theological or superstitious.[4] During the 18th and 19th century, research on epilepsy started advancing while, ancient physicians depended solely on clinical observation in order to determine the causes and symptoms of the condition. Thus, it took the step towards eliminating the stigma of superstitious and religious claims surrounding the cause of epilepsy. The second half of the 19th century, the treatment of epilepsy was based on delineation of pathophysiology and topographic localization of the epileptic seizures. There were several discoveries and experiments along the years which

led to the recognition of the existence of localized lesions on cortex of the brain that were involved in epileptic convulsions. Therefore, the following quotation: "Epilepsy is the name for occasional, sudden, excessive, rapid and local discharges of grey matter" became the definition of epilepsy in 1873 by John Hughling Jackson.[3]
As epilepsy is a prevailing neurological disorder where millions have been affected worldwide, the condition and its symptoms are being tackled through methods like epileptic seizure detection. The improvement of technology over the past years has escalated the field of doing detailed research. With increasing techniques and experimentations, we have managed to overcome the limitations from the past, thus narrowing the gap between several problems and solutions. Epileptic seizure detection means detecting the presence of EEG signals that are causing the seizures. Researchers have used different techniques to process the raw EEG signals and use different classifiers to detect seizure EEG signals. Detecting seizures beforehand prevents countless injuries[2] and rapidly promotes the therapies aimed at treating epilepsy, therefore increasing the importance of epileptic seizure detection. Although epileptic seizure detection is still a challenging issue but we have come a long way with the help of different technologies.

1.3 Aims and Objectives

The aim of our research is for the proposed model to achieve higher accuracy rate than the state-of the-art models. It has been seen that the International 10-20 system, which is a widely used electrode placement system, records electrical activity in different areas of the brain. For a patient not all the channel readings indicate seizure is occurring, as seizure can occur at any part of the brain at any time. So, there is no certainty that all the channels will give us information about the occurrence of seizure. If reduction of channel is not implemented while detecting seizure then that will not only cause computational delay but also hamper the overall accuracy of the detection. It is inevitable to ensure the use of the minimum number of channels possible to reduce system power consumption, a necessary step to maintain longer time of operation[5].
One of the challenging aspects of seizure detection is that different parts of the brain serve us with different signals. Not all of them are of similar nature. If we are unable to differentiate between channels then it makes it much more difficult for us to detect necessary signals for epilepsy. Among these multiple channels not all the sources are implying that a seizure is occurring. So, if readings of such channels are also included then there is a high chance that we will be unable to detect seizure accurately. In this regard, automatic channel selection is extremely important to focus only on the channels that consist of the most necessary information and ignore the unnecessary channels. The most significant information of the functional state of the human brain lies in five major brain waves distinguished by their frequency bands. These frequency bands are- delta band (0–4 Hz), theta band (3.5–7.5 Hz), alpha band (7.5–13 Hz), beta band (13–26 Hz), and gamma band (26–70 Hz)[6]. One of the main objectives of channel selection is to identify brain areas that generate class event activity. If all the channels are considered then there is a high chance that the false positive rate might increase drastically. But only using appropriate channels can not only increase the detection accuracy but also lower the false positive rate. Through our proposed model, we aim to attain exceeding precision in the results of the

epileptic detection than the existing state-of-the-art models. Results demonstrated in this paper have outperformed the existing state-of-the-art methods with 97.22%, 100% and 83.33% for 2 class classification and 91.67%, 91.67% and 80.56% for 4 class classification for the aforementioned classification techniques accordingly.
The objective of our system is to detect epileptic seizures with the data sets of patient's EEG signals as the input by observing if the seizure has some sort of pattern or if the feature extraction techniques can accurately detect different stages of epilepsy. At first, the data has to be preprocessed. The data originally found is a series of EEG readings from different electrodes set on different parts of the brain. The priority is to separate the electrode readings followed by extracting features from them. For the proposed model different feature classifications are used and then checked if they are able to detect the stages of seizures properly. So, in that case the original domain will be transformed and then features will be re-extracted to see if the model is now able to differentiate the various stages accurately. The features that are found in the feature extraction step are not enough to make a generalized seizure detection system. So, they are then re-transformed into wavelet domain, which results in higher detection rate. After extracting the relevant features, we need to do is classify them. The EEG signals recorded can be classified into two stages at first- (i) non seizure period (where there is no sign of epileptic seizure) and (ii) seizure period. The seizure period is then divided into four classes: (i) preictal period (30 minutes to 60 minutes prior to ictal period), (ii) ictal period (the actual seizure period), (iii) post-ictal period (30 minutes later than the ictal period) and (v) interictal period (the duration between post-ictal period to pre-ictal period). The classification is important to know which are ictal and which are inter-ictal and which are pre-ictal signals. For the proposed model classifiers such as, DTC, RFC and KNN classifiers are used. The seizure detection commences with the feature extraction and classification of the EEG signals.

1.4 Research Methodology

This model serves the purpose of differentiating different seizure stages. In the proposed model, feature extraction methods such as DFA, HFD, HJORTH parameters, Fisher, SVD have been used. Then they were transformed into Discrete Wavelet domain again. The proposed model produces two types of result i) Two stages – seizure and non-seizure and ii) four stages, which are- pre ictal, ictal, inter ictal and post ictal. Three different classifiers: DTC, RFC and KNN have been used to classify into seizure and non-seizure stages. The objective of our system is to detect epileptic seizures with the data sets of patient's EEG signals as the input.

1.5 Research Orientation

In section 2, we have exhibited the related works and researches that have been implemented by other researchers on epilepsy and seizure detection in the past. Then, section 3 describes the steps and methods we have used in order to bring about accurate results. Following that, section 4 demonstrates the implementation of our research through the execution of the research methodology. We have also illustrated the use of two dataset for clear representation and understanding for data classification. Then, section 5 shows the results and analysis achieved from the implementation of our methods. Lastly, in section 6 we have included the conclusion and future work of the proposed model of our research.

Chapter 2

Related Work

For detection and prediction of epileptic seizures, various methods can be used to interpret the EEG signals. One such method as done by Panda[7] et. al. involves the use of an SVM based classifier on EEGs. For the study, five types of EEG data were chosen which were divided in the following way: healthy subject with eye open condition, eye close condition, epileptic, and seizure signal from hippocampal area. Signals were preprocessed and decomposed up to the fifth level of the decomposition tree using the DWT. Various characteristics like energy, entropy, and standard deviation were calculated and used to classify signals. The results demonstrate that the detection of abnormal from normal EEG signals has a classification accuracy of about 91.2%. But it is not easy to detect seizures, because seizures can vary from person to person and also in terms of location. So a more generic approach towards detection is discussed in a paper by Parvez[8] et. al. regarding feature extraction of pre-ictal and interictal stages from EEGs using EMD along with DCT. This is followed by LS-SVM for classification purpose. This method has proven useful in terms of sensitivity, specificity and accuracy when it comes to classification of ictal and preictal stages. Lasefr[9] et. al. processed the EEG signal data for both time and frequency domain at first, and used Chebyshev filter to pre-process the data. Then Wavelet Analysis was used to decompose the signal into five sub-bands. This was done for both time and frequency domain. Then further processing was done using only the Delta sub-bands from here. Then feature extraction was done using DWT. A threshold was kept to find out what is the noisy part of the extracted signal. M.A. Hadj-Youcef[10] et.al. also conducted research in classifying EEG signals that focused on getting an accurate start to the process. It was done by decomposing the raw EEG signals with the help of DWT for extracting the maximum, minimum, range, standard deviation, energy, entropy. The decomposed signal was then reduced by PCA to further refine the data for a better representation of epileptic seizure data. After this SVM classifier was added to reach an accurate level of classification. An accuracy of 98% was reached. Parvez[2] and Paul[2] also worked with DCT-DWT, and SVD to be able to differentiate between ictal and interictal period. But the technique was applied on two different datasets, one big and one small (Epilepsy small dataset, 2012 and Epilepsy large dataset, 2013). For both the cases the accuracy was high. Through the process, it was also demonstrated different lobes of the brain show different phenomenon during ictal and interictal period. A different technique was applied by Liang[11] et. al, where the EEG features were extracted and was evaluated by the ApEn analysis and multiclass EEG were found. The ApEn analysis was further improved by including EEG sub-bands or autoregressive models. Then the PCA and GA on this was compared for selection of feature through different linear and non-linear methods so that it can be used for the next classification stage. In another study conducted by Ocak[12], normal and abnormal EEG signals went through fourth-level packet decomposition to create frequency bands. Then the ApnEn value from the decomposition tree were used to predict EEG data from previously collected frequency bands. Finally, GA was used to find optimal feature from here to maximize the performance of a LVQ-based normal and epileptic classifier. GA helped with increased accuracy which came out to be 94.3% for normal EEG epochs and 98% for epileptic EEG epochs. Recently, as stated by Pachori[13], applying a decomposition method called EMD has proven to be an efficient transformation technique. EEG signals are decomposed to IMF by EMD. Then a MF is calculated for each of the IMF found with the help of a FB expansion to differentiate between ictal and seizure-free intracranial EEG signals. Bajaj[14] et. al., also worked with differentiating between such signals so that classification between seizure and non-seizure EEG signals can be made to be accurate. EMD is used to generate IMFs and then the Hilbert Transformation of the IMFs were used for an analytic signal representation of IMFs. Two bandwidths are then calculated from here which is then classified using LS-SVM to differentiate between seizure and non-seizure EEG signals. When the process was paired with RBF, accuracy was 98.0%. When the process was paired with Morlet Kernel, the accuracy was 99.5%. Duman[15] et. al., focused on early prediction of seizures that focused on extracting the first 5 IMFs from EEG signals that were decomposed into IMFs. This method used the Hilbert-Huang Transform to create a result with an accuracy of 89.66%. On the other hand, Zhang[16] et. al. proposed a different technique which works with neural network as its basis. They devised four hand tasks that were complicated to perform and then formed the EEG signals created during the performing of the tasks. The obtained signals were pre-processed by decomposing them using wavelet transform. From this, a matrix was formed and SVD was applied and the singular value feature vectors were found from the matrix. These features were then used as an input to ANN that could differentiate between the complicated tasks. The classification accuracy obtained from this technique was 89%. To improve EEG classification accuracy, Dastidar[17] et. al. worked with three key components of wavelet-chaos-neural network, namely standard deviation (quantifying the signal variance), correlation dimension, and largest Lyapunov exponent (quantifying the non-linear chaotic dynamics of the signal). The study not only focused on improving accuracy but also on reducing computing time and output analysis by switching up the process through performing the study in two phases: band-specific analysis and mixed-band analysis. The process made it possible to achieve an accuracy as high as 96.7%. Bezobrazova[18] and Golovko[18] also used chaos theory, that is, they determined the largest Lyapunov's exponent. This was done by applying forecasting neural networks on the experimental results. Hybrid studies were also conducted to make the process more intelligent for diagnosis purposes. One such study conducted was by Polat[19] et. al.. In the study, a hybrid system was built using decision tree classifier and FFT to extract features that would help to detect epileptiform discharges in EEG. Steps such as the use of k-fold cross-validation, classification accuracy, and sensitivity and specificity values were taken to increase accuracy. 5- and 10-fold cross-validation resulted in 98.68% and 98.72% classification accuracies. There is also a method by Parvez[20] and Paul[20] involving the exploitation of spatiotemporal relationship of EEG signals where certain reference vectors are set in the EEG signals that can compare the change in state between ictal and preictal stages. This method is also less sensitive to artifacts which is crucial as artifacts can increase inaccuracy. This method proposed an accuracy of 91.95% with lower chances of false negatives. There are many more filtering and/or denoising techniques available. The challenges are faced due to variations found in brain signals. Nonetheless, several sampling, extraction, analysis and classification methods have been designed with increasing accuracy that makes this side of the research promising.

Chapter 3

Research Methodology

3.1 Our Methodology

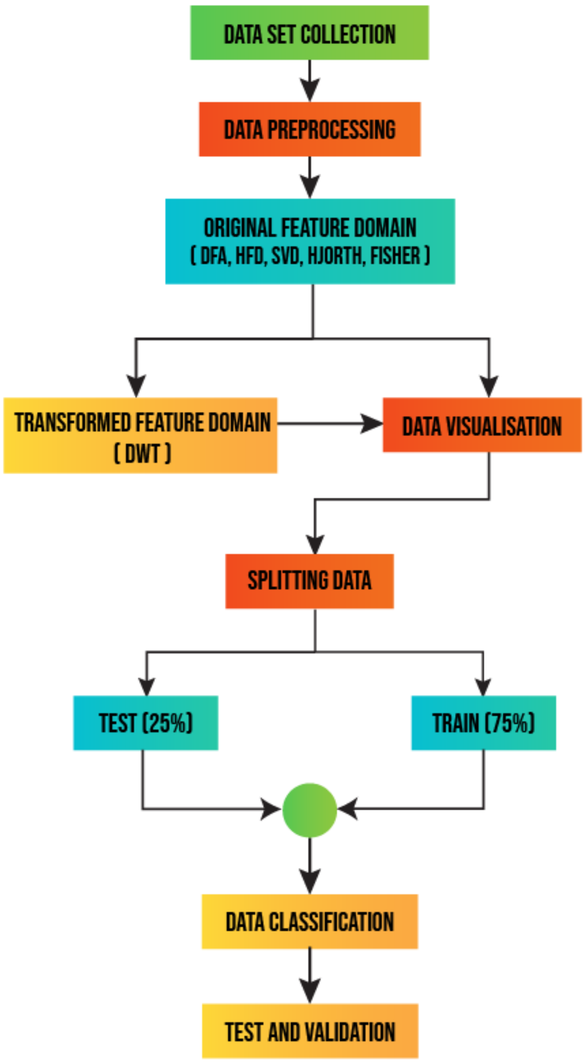

Figure 3.1: Work Flow Diagram

The work flow diagram in figure 3.1 depicts an overview of each and every step we have taken in order to train and evaluate our model. At first, we collected our required dataset. The gathered raw data was then cleaned, formatted, and organized using data preprocessing. Following that, we utilized different feature extraction approaches to extract the characteristics of the EEG signals in the original domain, such as DFA, HFD, HJORTH parameters, Fisher and SVD. Then, the original domain of the extracted features was transformed using DWT. After that we did data visualisation on both the original feature domain and transformed feature domain using t-SNE and statistical testing. Next, we trained 75% of our data and left the remaining 25% for testing. Then, for both the trained and testing data, we utilized data classifiers (DTC, RFC, and KNN) in both the original domain and transformed domain to test and validate our detection approach.

3.2 Original Feature Domain

In order to extract features from EEG signals, we utilized several feature extraction methods such as DFA, HFD, HJORTH parameters, Fisher, and SVD. All of these feature extraction techniques are aiding us in estimating the frequency of the seizures.

3.2.1 Detrended Fluctuation Analysis (DFA)

DFA is one of the main methods of analyzing temporal structure in scale free dynamic. DFA is a feature extraction technique which is widely used for analysis of non-stationary time series which has been applied to EEG signals[21]. The DFA technique can be successfully used to both resting-state and task-dependent recordings. It can also be utilized for measuring brain activity during diverse activities, such as mental counting, visual and motor imagery, or even during presentation of different stimuli[22]. The advantages of DFA over other approaches include the ability to discover inherent self-similarity buried in a seemingly non-stationary time series and the avoidance of erroneous detection of apparent self-similarity, which may be an artifact of external trends[23].
In order to perform DFA, at first, we need to convert the signal into mean centered cumulative sum. Then we have to define a number of scales that are logarithmically spaced. Because these scales get quite lengthy, it is not a technique to be applied on brief task related epics where data epochs are only a few seconds long. We need longer time series to do this analysis. Following that, we have to split the data based on the scales, detrend each segment, and compute the RMS for each of them. We need to repeat this for each scale. Lastly, we have to compute the linear fit between the log-scales and log-RMS. To summarize, in order to execute DFA[24], given that there are specified time series, we compute X_t (mean-centered cumulative sum) using the following formula:

$$X_t = \sum_{i=1}^{t} (x_i - \langle x \rangle)$$

Where, x_i is the bounded time series, $<x>$ is the mean value for time series. Then we calculate the fluctuation using the following formula:

$$F(n) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(X_t - Y_t)^2}$$

A single row from a file containing the electrode data has been converted into a one-dimensional array and DFA has been applied on the array. Since DFA analyzes the fluctuation of the data that are on the same column, a 1D array has been sent as input and a single scalar value has been returned.

3.2.2 Higuchi Fractal Dimension (HFD)

FD is a fast non-linear measure of the complexity of a signal. Its reduction is noticed because of increased EEG synchrony. This has been seen in epileptic convulsions, which show enhanced synchronous oscillations and decreased arrhythmic activity. HFD is one of the most effective techniques for determining a time series' FD. It can obtain the fractal dimension of a time series signal even when just a few data points are provided[25]. When compared to other techniques, it provides a more accurate measure of signal complexity.
Given a time series X:1,. . . ,N→R with data points and a parameter $k_{max} \geq 2$ the HFD of is calculated for each $x \in \{1, \ldots, k\}$ and $m \in \{1, \ldots, k\}$ define the length $L_m(k)$ as,

$$L_m(k) = \frac{N-1}{\left[\frac{N-m}{k}\right]k^2} \sum_{i=1}^{\left[\frac{N-m}{k}\right]} |X_n(m+ik) - X_N(m+(i-1)k)|$$

The length (k) is defined by the mean value of the k lengths $L_1(k), \ldots, L_k(k)$,

$$L(k) = \frac{1}{k}\sum_{m=1}^{k} L_m(k) \mid$$

The HFD of the time-series X is defined as the slope of the best-fitting linear function over the data points $\left\{\left(\log\frac{1}{k}, \log L(k)\right)\right\}$
The row of data has been converted into a 1D array and sent as time series and number of iterations (constant K) has been set to 8 for HFD. HFD returns a scalar value for each time series[26].

3.2.3 HJORTH Parameters

The HJORTH parameter is a method of showing statistical properties of a signal and it has three types of parameters- Activity, Mobility, and Complexity. The vari-

ance of the time function is known as the activity parameter which can represent the surface of the power spectrum in frequency domain. Here the value of activity yields a large or small number depending on how many or few high frequency components of the signal are present. The mobility parameter is defined as the square root of the variance ratio between the signal's first derivative and the signal itself. This value is proportional to the standard deviation of the power spectrum. As the form of the signal approaches that of a pure sine wave, the value of Complexity approaches one. These three parameters contain information about frequency spectrum of a signal[27].
Among these parameters we worked with the HJORTH mobility parameter[28] as a part of feature extraction,

$$\text{Mobility } = \sqrt{\frac{\operatorname{var}\left(\frac{dy(t)}{dt}\right)}{\operatorname{var}(y(t))}}$$

A 1D array is needed as input for HJORTH. But what makes the HJORTH parameter different from the other feature extractors is that HJORTH returns a tuple containing 2 values as output which are mobility and complexity. Here only mobility was used as a feature. The outputs that HJORTH generates are both scalars.

3.2.4 Fisher

Using the scalar feature for Fisher analysis[29], we can define Fisher information in the following way:

$$I = \sum_{i=1}^{M-1} \frac{\left(\bar{\sigma}_{i+1} - \bar{\sigma}_i\right)^2}{\bar{\sigma}_l}$$

This method is widely used to extract features by keeping a constant which is called Tau (τ). Tau (τ) refers to the lag. Moreover, another value is passed in the parameter which is DE. Both the values were passed as 4 along with a 1D array for each row in the proposed model. This produces a scalar value for each row as well.

3.2.5 Singular Value Decomposition (SVD)

SVD technique helps us to reduce bulks of data into key features that are necessary to analyze and describe that large data further. The SVD identifies the smallest number of dimensions necessary to express a matrix or linear transformation. Due to data redundancy, multidimensional data is frequently represented equivalently in fewer dimensions. If an n-dimensional vector set all lies in the same k-dimensional subspace, $k < n$, then each n-vector has only k degrees of freedom and can be uniquely represented by k integers. Because of the inherent connections that exist in nature, the SVD is an excellent choice for feature reduction. The magnitudes of the singular values revealed by the SVD will indicate if no further reduction is

possible[30].
For SVD, let the input signal be $[x_1, x_2, \ldots, x_N]$. Thus, the delay vector is constructed as[29],

$$y(i) = \left[x_i, x_{i+r}, \ldots, x_{i+(d_E-1)r}\right]$$

where τ is the delay and d_E is the embedding dimension. We considered, d_E =2 and τ= 2. The embedding space is then constructed by –

$$Y = \left[y(1), y(2), \ldots, y\left(N - \left(d_E - 1\right)\tau\right)\right]^T$$

The SVD is then performed on matrix y to produce M singular values, $\sigma_1, \ldots, \sigma_m$, known as the singular spectrum.
The SVD entropy is then defined as-

$$H_{SVD} = -\sum_{i=1}^{M} \bar{\sigma}_l \log_2 \bar{\sigma}_l$$

where M is the number of singular values and $\overline{\sigma_l}, \ldots \overline{\sigma_M}$ are normalized singular values such that $\overline{\sigma_l} = \frac{\sigma_i}{\sum_{j=1}^{M} \sigma_j}$. SVD entropy is a scalar feature[29].

3.3 Transformed Feature Domain

Studies have shown that dimensionality reduction in the feature set can yield better results than the original domain[31], For this reason, the features extracted from DFA, HFD, Hjorth, Fisher and SVD are again decomposed into multiple scales using DWT.

3.3.1 Discrete Wavelet Transform (DWT)

A wavelet is a time-localized wavelike oscillation. DWT employs a limited collection of wavelets, which are specified at certain scales and locations. This transform decomposes the signal into a set of wavelets that are mutually orthogonal.
While analyzing EEG signal, DWT is one of the most widely used techniques because of its resolution of frequency and time. DWT can be represented by an orthogonal matrix W of size n×n[32]. Let, x be a vector, its length is n. WT of vector x is vector y. Here, $y = Wx$.
But instead of multiplying x by the matrix W to figure out vector y, a more efficient approach is to use Mallat's pyramid algorithm to compute the transform. This algorithm's stages are as follows: expanding the input vector, filtering it using particular low-pass and high-pass filters, cropping the center section of the results,

and decimating the results. The low-pass branch coefficients are referred to as "approximations," whereas the high-pass branch coefficients are referred to as "details." This single transformation step can be repeated with the approximations serving as the input. Transformation depth is defined as the number of such repetitions[32].

3.4 Data Visualisation

We utilized two visualization approaches – t-SNE and statistical testing, to determine how much better the transformed feature domain is than the original feature domain and whether or not our model is patient dependent.

3.4.1 t-Distributed Stochastic Neighbor Embedding (t-SNE)

The t-SNE approach is an unsupervised, non-linear technique that is mostly used for data exploration and visualization of high-dimensional data. Only small pairwise distances or local similarities are preserved by t-SNE. It has been used to visualize the data which cannot be separated by a straight line. t-SNE is an iterative statistical process which visualizes high-dimensional data by giving each datapoint a location in a two or three-dimensional map.
The t-SNE algorithm has two main stages. First, t-SNE creates a probability distribution from the data where similar data are given a higher probability and dissimilar data are given a lower probability[33].

$$p_{j|i} = \frac{\exp\left(-\frac{\|x_i - x_j\|^2}{2{\sigma_i}^2}\right)}{\sum_{k \neq i} \exp\left(-\frac{\|x_i - x_k\|^2}{2{\sigma_i}^2}\right)}$$

A value of 30 is used here for perplexity. Perplexity is the target number of neighbors for our central point. Higher perplexity causes higher value variance. The typical range for perplexity is from 5 to 50. The used perplexity value is then used with the value of sigma that produces probability distribution using a binary search.
Second, t-SNE creates a low-dimensional space with the same number of points as in the original space. Which minimizes the KL divergence between the two distributions with respect to the locations of the points in the map[33].

$$KL(P||Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}}$$

t-SNE has been used to plot data from different patients to study and compare the data. It has also been used to further study different features of a patient and compare between them.

3.4.2 Statistical Testing

P-value has been used to see how the features are connected to each other and how they co-relate. It can also be used to compare the distribution of the features. P-values shows how incompatible the data are with a specific statistical model (usually with a null-hypothesis). A p-value of less than 0.05 is a cutoff point for statistical significance.[34]

Both patient wise data and multiple patient data have been used to measure the p-value between seizure and non-seizure states. The pairs taken for p-value are pre-ictal and ictal, ictal and post-ictal pre-ictal and post-ictal. We have used Mann-Whitney U test for calculating p values.

3.5 Classifiers

We used three different classifiers- DTC, RFC, and KNN- to assess the accuracy of our detection approach in both two (seizure and non-seizure) and four (seizure and class (ictal, preictal, interictal and postictal).

3.5.1 Decision Tree Classifier (DTC)

DTC is a widely used classifier in classical machine learning approach. It basically creates a tree structure to predict and make decisions based on the branches. Branches are created using the outcomes and the leaf nodes are used as classes for prediction[35]. Entropy has been used as the criterion in DTC.

3.5.2 Random Forest Classifier (RFC)

RFC works in a similar approach to DTC. RFC creates multiple decision trees and each tree gives us a class prediction. The class with the most prediction is then used as the primary prediction. RFC is a simple but powerful approach as it uses the concept of committee in making the decision[36].

3.5.3 KNN Classifier

KNN classification is a simple non-parametric classification method. It identifies the nearest neighbors and using those neighbors, predict the class. It is a low efficient approach but an effective one[37]. KNN classifier also takes another parameter K along with the data. The K refers to the number of neighbors to take when classifying. K has been set to 3 for this case.

Chapter 4

Implementation

4.1 Dataset

4.1.1 Source

The Children's Hospital Boston Dataset, also known as the 'CHB-MIT Scalp EEG Database'[38], [39].

4.1.2 Dataset Description

The dataset recorded at the Children's Hospital Boston is one of the most widely used dataset that consists of EEG recordings of 23 pediatric participants who happen to have medically intractable seizures. All the recordings are grouped into 23 cases and were collected from 22 participants initially (5 males, ages 3–22; and 17 females, ages 1.5–19). Later on in December 2010, another participant's diagnosis was added which makes a total of 23 patients in this dataset. In the dataset, one of the female participants has been examined twice (chb21 and chb01). Each case (chb01, chb02, and so on) contains 9 to 42 continuous .edf files from a single subject. Due to hardware restrictions, there are pauses between sequentially numbered .edf files during which the signals are not captured; in most situations, the intervals are shorter than 10 seconds, but there are times when the gaps are significantly longer. The .edf files in most cases include exactly one hour of digitized EEG signals, while those in case chb10 are two hours long, and those in cases chb04, chb06, chb07, chb09, and chb23 are four hours long.

For each of the patients the data sampling rate is 256 Hz with 16-bit resolution and there are mostly 23 channels in the edf files. For these EEG recordings the international 10-20 system of electrode positions and nomenclature had been used. The file record contains a total of 664 edf files, among those files 129 files contain one or more seizures. There are a total of 198 seizures in all these files. We are dividing the epileptic EEG signals in four stages (preictal, ictal, interictal, post ictal) for our experimental purpose. The preictal phase refers to the time before the seizure, ictal phase is the time from the start to the end of the seizure activity, interictal phase indicates the period between seizures and post ictal phase refers to the immediate period after the seizure ends.

Other signals, such as an ECG signal in the final 36 files of case chb04 and a VNS signal in the final 18 files of case chb09, are captured in a few records. To provide an easy-to-read display style, up to 5 "dummy" signals (labeled "-") were inserted among the EEG signals in certain situations; these dummy signals can be disregarded. The tabular data of the dataset can be found in Table- 4.1.

Total number of patients	23	Total number of seizures	198
Age range	1.5-22	Sampling Rate	256 Hz
Total edf files	664	Duration of files	1/2/4 hours
Files with seizure	129	channels per edf file	23-28

Table 4.1: Summary of the CHB-MIT Scalp EEG Database

As mentioned earlier, we have used the international 10-20 system of electrode positions which is used to record the voluntary EEG. This system helps us to record the electric potentials on the scalp by using electrodes[40] that we can see in Figure-4.1. These electrodes are located on the surface of the scalp.

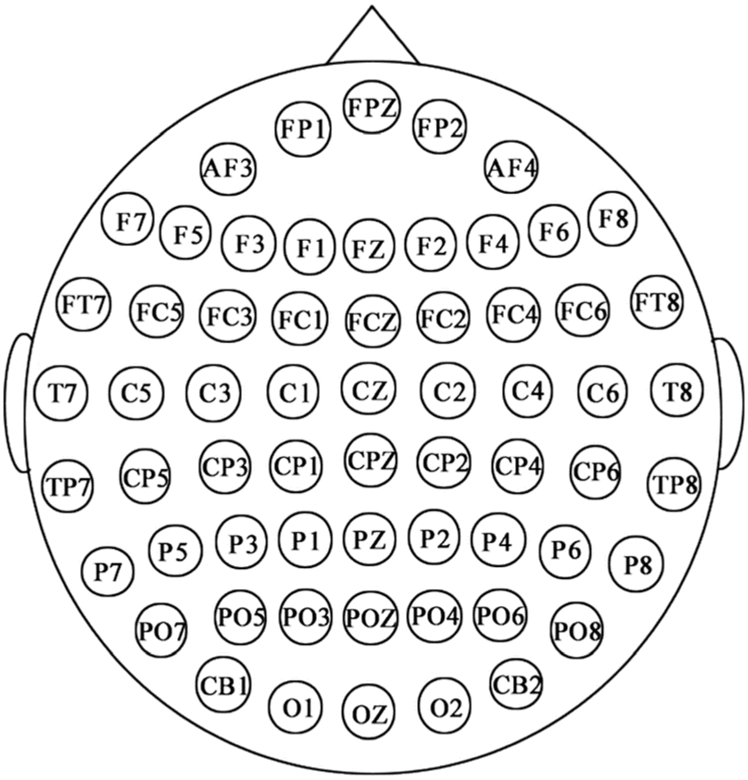

Figure 4.1: The international 10-20 system of electrode position[12]

4.1.3 Data Sample

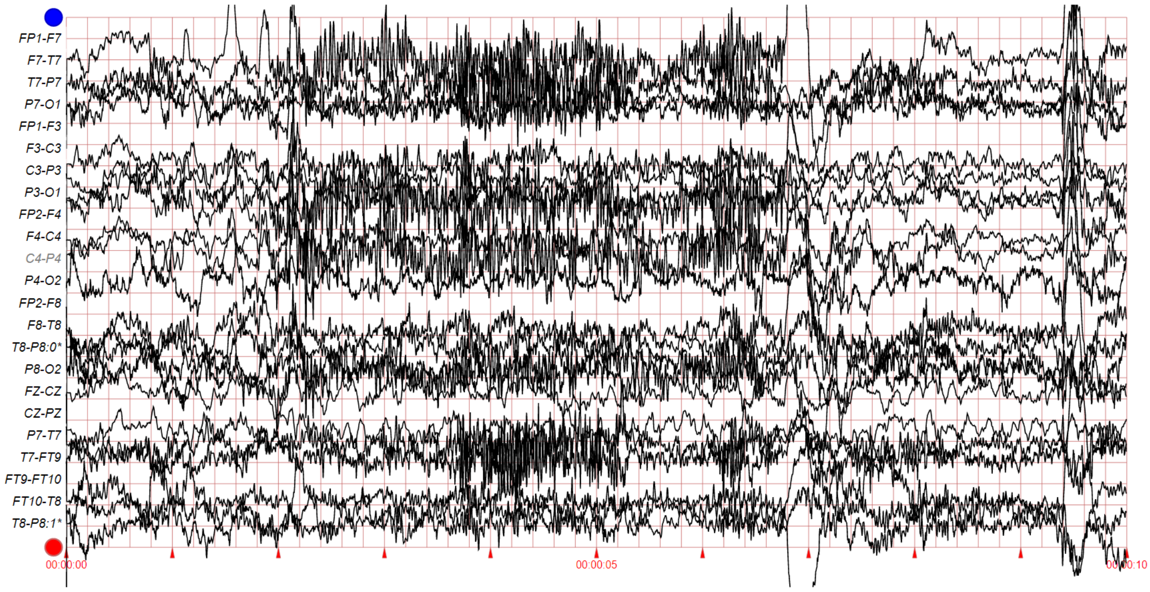

Figure 4.2: Sample Data of Non-Seizure Signals of the Dataset

In figure 4.2, we can see a ten second sample data from file 1 of patient 1. This sample shows the non-seizure readings from the EEG signal.

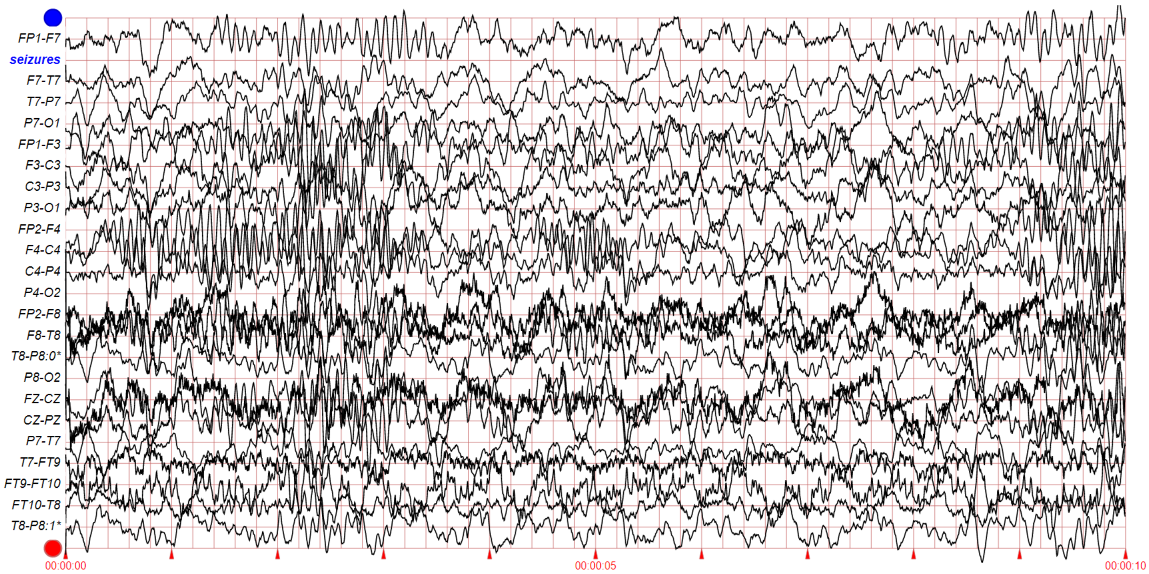

Figure 4.3: Sample Data of Seizure Signals of the Dataset

In figure 4.3, we can see a ten second sample data from file 3 of patient 1. This sample shows the seizure readings from the EEG signal.
In both cases, the Y-axis represents the different electrode sources and the X-axis represents time.

4.2 Data Pre-processing

The data sampling rate for each patient is 256 Hz. The number of electrodes and amount of time a patient has been examined varies from patient to patient. The .edf files are converted to .csv files where the rows in the initial csv file of each patient refers to the total number of samples (256 Hz) and the column indicates the number of electrodes used for that patient's examination. As per second we get 256 samples, so in one hour we are getting a total of 921600 samples (when the reading is taken for one hour for a particular patient). So, the total row count in the initial .csv file will be 921600.

	# FP1-F7	F7-T7	T7-P7	....	FT10-T8	T8-P8-1
1	-1.78E-05	3.93E-05	-3.71E-06	....	-3.97E-05	-6.00E-05
2	1.95E-07	1.95E-07	1.95E-07	....	1.95E-07	1.95E-07
3	1.95E-07	1.95E-07	1.95E-07	....	1.95E-07	1.95E-07
4	5.86E-07	1.95E-07	1.95E-07	....	-1.95E-07	1.95E-07
5	1.95E-07	1.95E-07	1.95E-07	....	1.95E-07	1.95E-07
6	-1.37E-06	1.76E-06	1.95E-07	....	2.54E-06	1.37E-06
....	....	....	....	....	....	....
921599	-8.79E-06	-2.25E-05	6.04E-05	....	3.85E-05	2.64E-05
921600	-4.88E-06	-2.05E-05	5.92E-05	....	3.61E-05	3.22E-05

Table 4.2: Sample of the converted .csv file

Secondly, the channels were separated as individual files. There are 21 general electrodes that are used to take EEG readings and other virtual electrodes can also be used. There are approximately around 90-95 channels used in this dataset to take patient EEG reading. By analyzing all the seizure files, we get to see that most of the seizure periods were around 45-60 seconds. So, we took a 60 seconds time frame to divide the entire one hour reading of a patient. Here, most of the files are converted into a file of 60 windows of each containing 60 seconds of data in each row. By doing this the resulting input data will be a 60*15360 matrix for each electrode from each electrode of each file. Thus, by preprocessing the raw dataset from the patient files, we formed a more efficient dataset which will help with feature extraction and further analysis.

	1	2	3	....	15359	15360
1	-1.78E-05	1.95E-07	1.95E-07	....	-1.58E-05	-1.74E-05
2	-1.66E-05	-8.79E-06	-3.32E-06	....	2.87E-05	3.03E-05
3	2.68E-05	1.89E-05	1.54E-05	....	8.01E-06	1.23E-05
4	1.74E-05	2.01E-05	1.66E-05	....	4.12E-05	4.28E-05
....	....	....	....	....	....	....
59	2.75E-05	2.48E-05	2.13E-05	....	8.79E-06	2.09E-05
60	2.36E-05	2.21E-05	2.40E-05	....	-8.79E-06	-4.88E-06

Table 4.3: Sample of the #FP1-F7 electrode readings split into 60 seconds time windows

4.3 Feature Extraction

DFA, HFD, Fisher, SVD and HJORTH are used to extract features from the time windows from each electrode. For DFA, a single row from a file containing the electrode data has been converted into a 1D array and DFA has been applied on the array. Since DFA analyzes the fluctuation of the data that are on the same column, a 1D array has been sent as input and a single scalar value has been returned. We have also sent each row of data as 1D array for HJORTH feature extraction. But HJORTH returns a tuple containing 2 values as output which are mobility and complexity. Here only mobility was used as a feature. The outputs that HJORTH generates are both scalars. For HFD, the row of data has been converted into a 1D array and sent as time series and number of iterations (constant K) has been set to 8. HFD returns a scalar value for each time series which is used as the feature here. Fisher is widely used to extract features by keeping a constant which is called Tau (τ). Here, Tau (τ) refers to the lag. Moreover, another value is passed in the parameter which is DE. Both the values were passed as 4 along with a 1D array for each row in the proposed model. This produces a scalar value for each row as well. SVD also requires τ which is the delay and D_E which is the embedding dimension in order to extract feature. Here, 2 is used for both cases. Each file contains the features from the data of a single electrode and their corresponding 4 and 2 class targeting which we have added using the information given along with the dataset.

DFA	HFD	FISHER	SVD	HJORTH	4 class	2 class
0.6139	0.1850	0.9525	0.3189	0.0009	pre ictal	non-seizure
0.6202	0.2204	0.9660	0.3254	0.0010	pre ictal	non-seizure
0.3873	0.0701	0.7938	0.2556	0.0006	ictal	seizure
0.5531	0.1092	0.8047	0.2555	0.0006	ictal	seizure
0.9420	0.2902	0.9340	0.3314	0.0011	inter ictal	seizure

Table 4.4: Dataset using features extracted using DFA, HFD, Fisher, SVD and HJORTH

4.4 Feature Domain Transform

The features extracted from DFA, HFD, HJORTH, Fisher and SVD are again decomposed here into multiple scales using DWT. The multiple scales are then passed through a high pass filter which gives detailed information and a low pass filter which gives the approximation co-efficient[41]. Taking only the approximation co-efficient value gives us a reduced dimension of 3 for the features which is used to further study the classification of different stages.

DWT-1	DWT-2	DWT-3	4 class	2 class
0.832372	0.534052	0.89908	pre ictal	non-seizure
0.692854	0.556998	0.91325	pre ictal	non-seizure
0.341367	0.461581	0.742154	ictal	seizure
0.639327	0.486255	0.749723	ictal	seizure
1.098111	0.606554	0.894879	inter ictal	seizure
1.088935	0.698754	1.108831	inter ictal	seizure
0.72514	0.558431	0.852786	inter ictal	seizure

Table 4.5: Dataset After Feature Domain is Transformed

4.5 Machine Learning Based Classification

We have used DTC, RFC and KNN classifiers for the classification of data. The data has been split into 75:25 ratio for training and testing the models. The classification has been firstly done on the 5 features extracted from DFA, HFD, HJORTH, Fisher and SVD. Later, the reduced dimension from DWT has been used for classification. The 2 different feature extraction methods give us an opportunity to study the prediction differences. For both data models, the classification has been made in 2 and 4 classes (2 having seizure and non-seizure states and 4 having Pre-Ictal, Ictal, Inter-Ictal and Post-Ictal states) in order to further make a comparison of the prediction of the models.

Chapter 5

Results and Analysis

5.1 Statistical Relationship

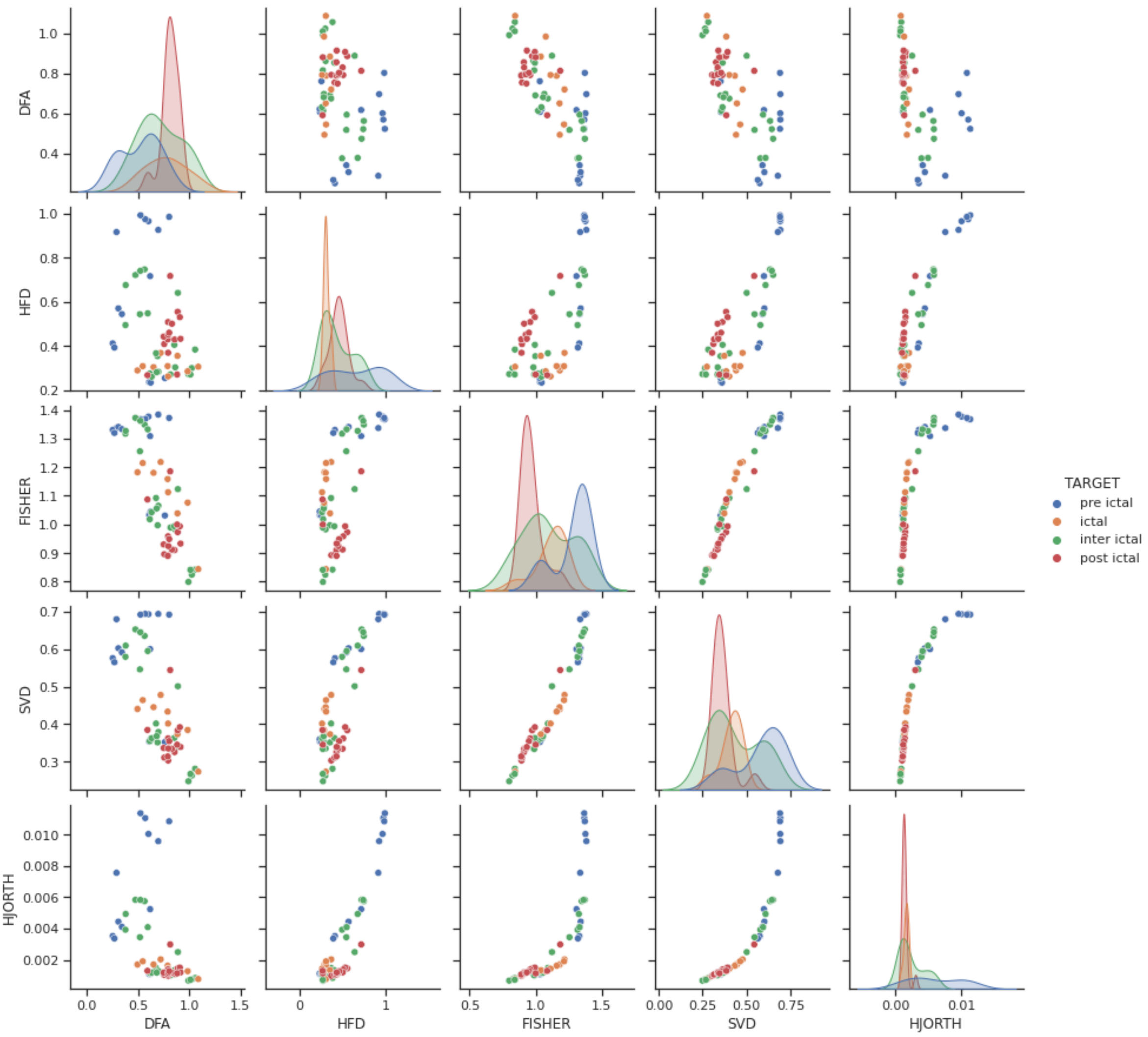

Figure 5.1: Bivariate relationships between combinations of features extracted (Single Patient)

In figure-5.1 we can see that it is a n*n matrix. Here, n is the features extracted. The diagonal plots are specific features in their own domain. For example, in Fisher vs Fisher domain we can see different peaks for different seizure stages. This statistical relationship was conducted on a single patients file. The features extracted tells us that there is some sort of correlation amongst them. If we notice one of the horizontal plots, in SVD vs Fisher there are clusters which may indicate that SVD and Fisher are one of the visible features that can be used.

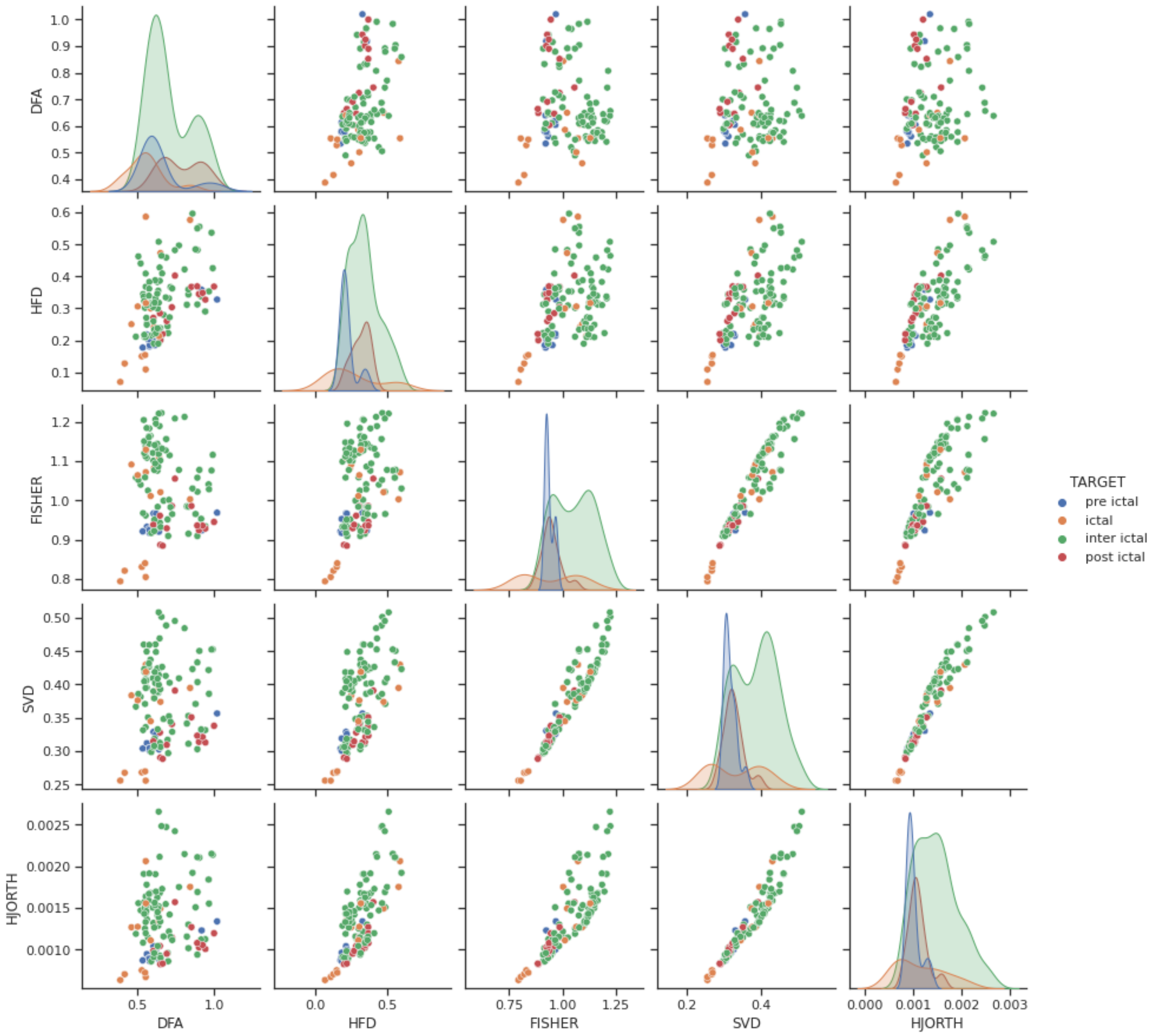

Figure 5.2: Bivariate relationships between combinations of features extracted (Multiple Patient)

While extracting the features for SVD parameters we passed a 1D array and we also passed the value of Tau and DE as 2 in the parameter. And for fisher the value of Tau and DE was passed as 4. In figure-5.2 This is a feature plot for multiple patients. Here we can see that there are no features that can clearly differentiate amongst the seizure stages. The parameter values for SVD and Fisher feature extraction techniques were also the same as before.

5.2 Data Visualisation

In this section there are multiple t-SNE plots to show the correlation ship among the extracted feature.

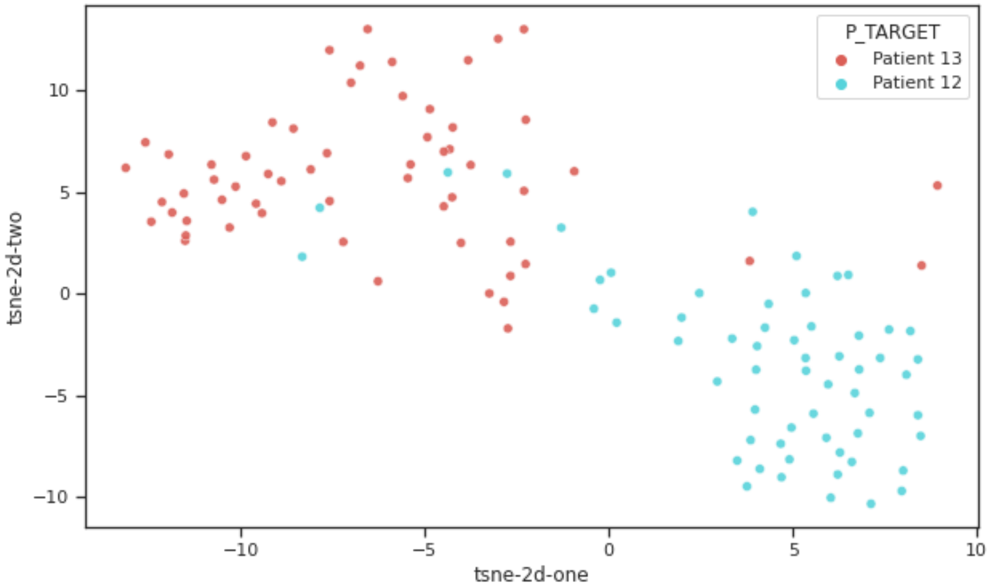

Figure 5.3: t-SNE plot of two different subjects (Feature Cluster)

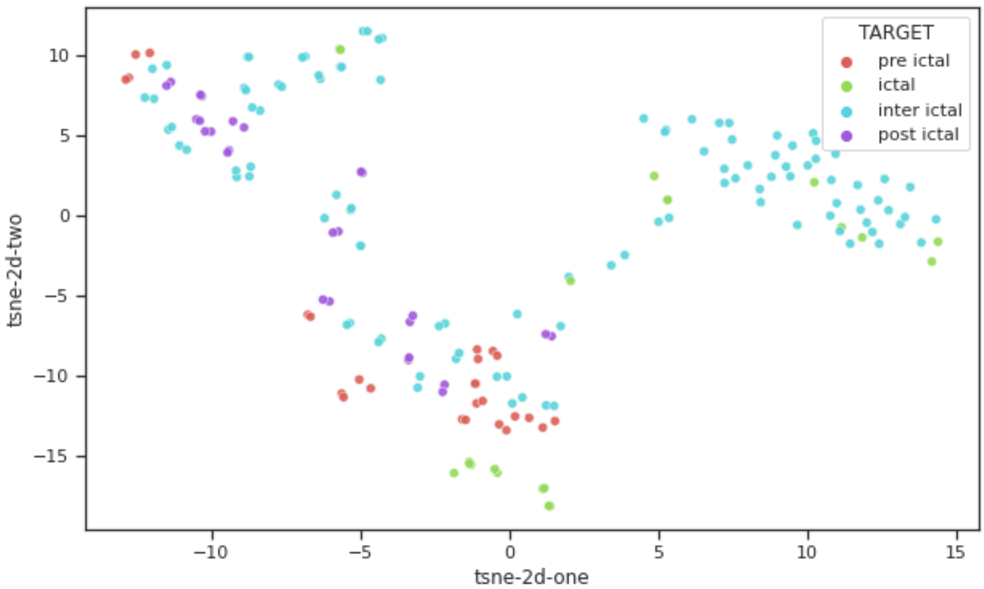

Figure 5.4: t-SNE plot of two different subjects (Seizure stages)

In figure- 5.3 we can see a t-SNE plot of multiple subjects in the original feature domain. This is to show that the features extracted from the same electrode of two different patients show different readings. We can see two different clusters in the figure which convey the fact the features extracted are patient based and this feature extraction-based model is a subject dependent model. All the seizure and non-seizure stages for the patients are going to be in different range. So, as a result the model will fail to detect seizure and non-seizure stages for different subjects.

The t-SNE plot in figure- 5.4 shows t-SNE plot of one patient in the original feature domain. As it can be seen that there are a few clusters which can indicate the seizure and non-seizure areas. So, it can be further confirmed that the features extraction techniques used previously are subject dependent. The model in the original feature domain can detect seizure and non-seizure areas.

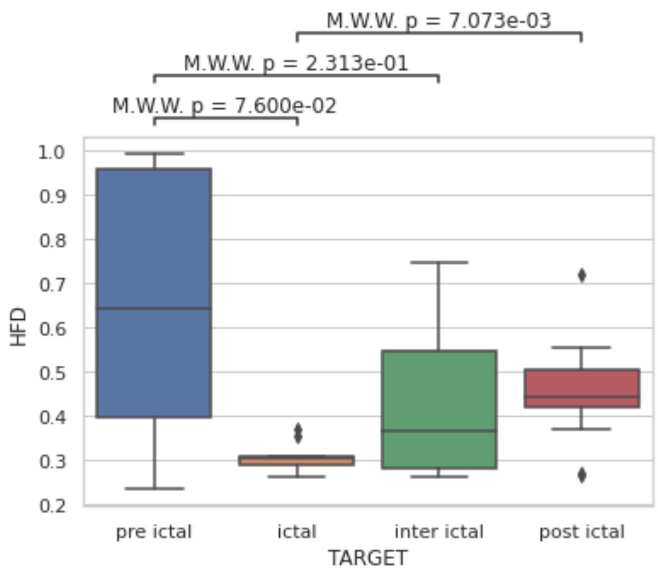

Figure 5.5: Whisker box plot for HFD (single patient)

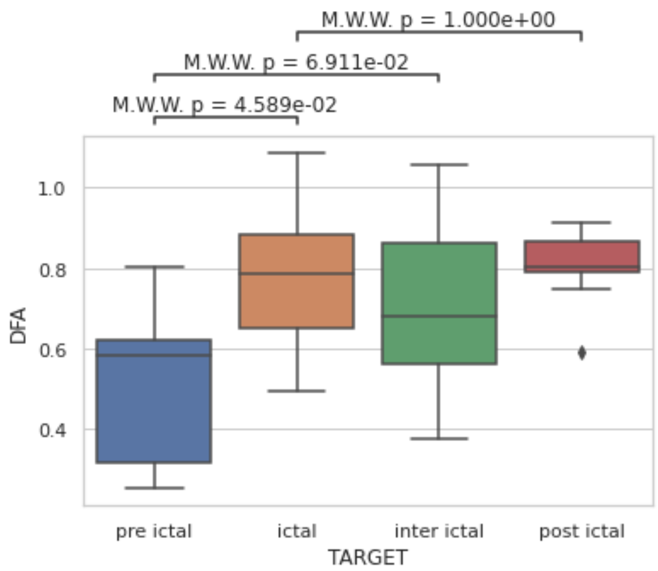

Figure 5.6: Whisker box plot for DFA (single patient)

In figure- 5.5 the p value test was conducted on the extracted features to prove the significance of the seizure stages. If the value of p is less than 0.05 it is considered that the distribution is significant. For p values calculation Mann-Whitney U test was conducted on the single patient's file.
As per the figures- 5.5-5.9 it can be seen the p values of seizure and non-seizure are of significant distribution. For example, in Fisher domain we can see the distribution

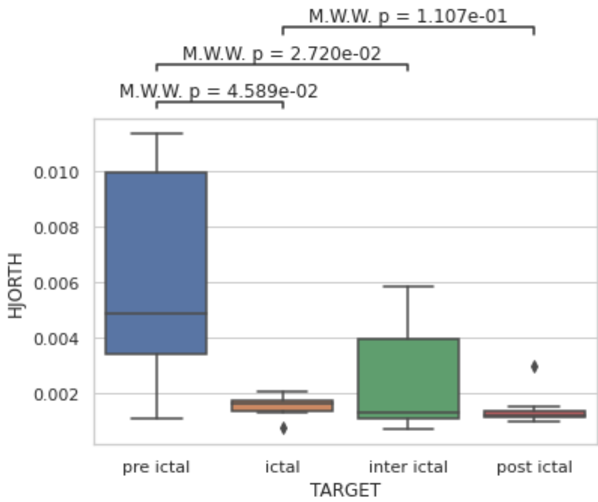

Figure 5.7: Whisker box plot for HJORTH (single patient)

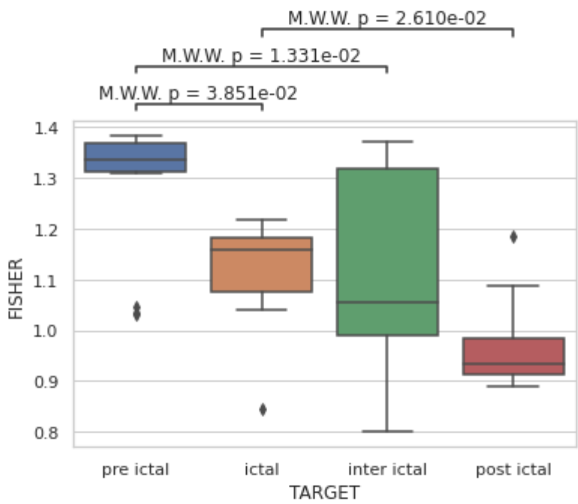

Figure 5.8: Whisker box plot for Fisher (single patient)

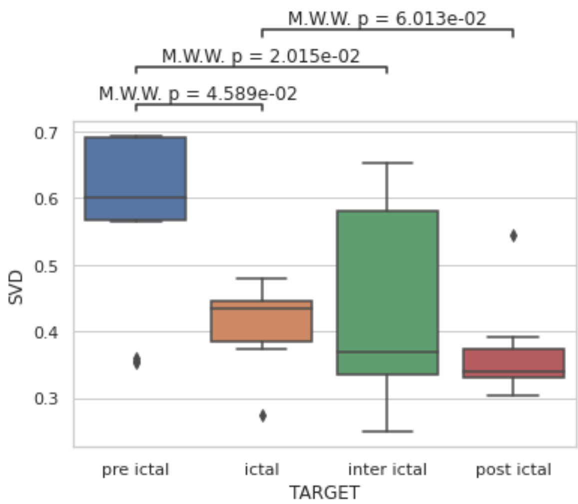

Figure 5.9: Whisker box plot for SVD (single patient)

of ictal and pre ictal and post ictal are significant as the values are well below 0.05. This goes on to further prove the case that the feature extractors used in this case are all patient or subject dependent.

5.3 Significance of Transformed Domain

So, in order to perform a better classification of seizure and non-seizure stages the extracted features were the transformed into DWT. DWT produces two values which are approximate coefficient and details coefficient. The approximate coefficient has been considered as features in this case.

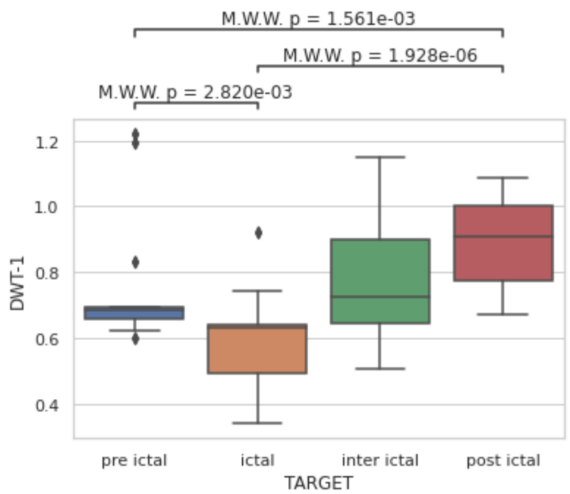

Figure 5.10: Whisker box plot for transformed feature domain (Multiple patient: DWT-1)

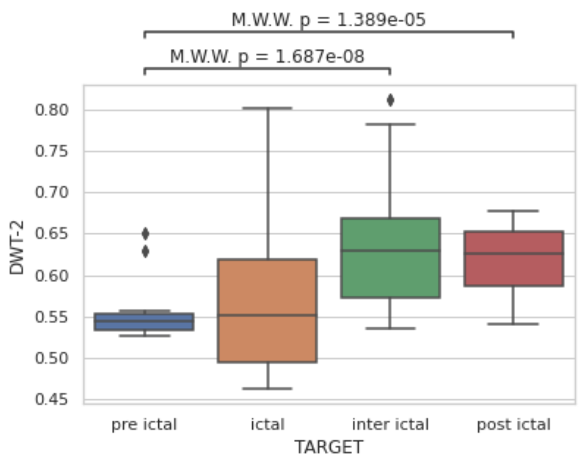

Figure 5.11: Whisker box plot for transformed feature domain (Multiple patient: DWT-2)

The figure-5.10 and 5.11 shows the p value significance of the re-extracted features from the previous 5 features. Which were- HFD, SVD, HJORTH, DFA and Fisher. The domain transform technique provides two features which can be named as DWT-1 and DWT-2. The above figure shows that the newly extracted features have been able to distinguish amongst the 4 stages of seizures with a p value of less than 0.05. In p test value, a value less than 0.05 between a pair means that the distribution is significant.

5.4 Classification

5.4.1 Classification using Original Domain Features

For the proposed model three different type of classifiers have been used, which are- DTC, RFC and KNN classifier. In the original feature domain 5 feature extraction techniques were used. In this section two types of classification have been attempted, which are- 2 class and 4 class. 2 class classification were set as seizure and non-seizure and 4 class classification was set as pre ictal, ictal, inter ictal and post ictal. This feature classification was based on multiple patients who has multiple seizure occurrences. The patients who were selected are patient-6, 9, 10, 12, 15, 17 and 22. The classification performance of the aforementioned classifier techniques on the original feature extraction domain are given in table-5.1 and table-5.2. The tabular data shows the classification accuracy of the classifiers in original domain.

Single Subject (Original Domain)		
Classifier	2 Class (accuracy)	4 Class (accuracy)
Decision Tree Classifier	75.0%	66.67%
Random Forrest Classifier	91.0%	75.0%
KNN Classifier	91.67%	78.72%

Table 5.1: Single Subject Classification in Original Feature Domain

Multiple Subject (Original Domain)		
Classifier	2 Class (accuracy)	4 Class (accuracy)
Decision Tree Classifier	86.11%	80.55%
Random Forrest Classifier	91.67%	86.11%
KNN Classifier	63.89%	61.11%

Table 5.2: Multiple Subject Classification in Original Feature Domain

5.4.2 Classification using Transformed Domain Features

As it can be said that the feature extraction methods that were used have been subject dependent, this model proposes a novel method to extract features from the previously extracted features using DWT. The classifiers that were used previously have also been used in this section. The achieved classification results are given in table-5.3. The tabular data shows the classification accuracy of the classifiers in transformed domain. The test subjects in this case are also the same as the cases in original domain.

Single Subject (Transformed Domain)		
Classifier	2 Class (accuracy)	4 Class (accuracy)
Decision Tree Classifier	91.67%	66.67%
Random Forrest Classifier	83.33%	58.33%
KNN Classifier	87.67%	66.67%

Table 5.3: Single Subject Classification in Transformed Feature Domain

Multiple Subject (Transformed Domain)		
Classifier	2 Class (accuracy)	4 Class (accuracy)
Decision Tree Classifier	97.22%	91.67%
Random Forrest Classifier	100%	91.67%
KNN Classifier	83.33%	80.56%

Table 5.4: Multiple Subject Classification in Transformed Feature Domain

It can be seen in table- 5.3 that all the classifier has performed better in the transformed domain compared to the original domain. KNN classifier has achieved visibly higher accuracy compared to what it achieved in the original domain. But in table-5.4 it can be seen that the features for single patient is not effective while detecting 4 classes. So, it can be said that DWT transformed features are effectively better for multiple test subjects.

Chapter 6

Conclusion

In this paper, we have proposed a novel method for detecting epileptic seizures. We have used different feature extraction methods like - DFA, HFD, HJORTH parameters, Fisher and SVD in order to extract the features from the EEG signals. All of these feature extraction approaches are assisting us in determining the frequency of the seizures. Following that, the original domain of the extracted features are transformed using DWT. DWT is employed to transfer the signal in the time domain[42]. By doing this transformation we are able to accurately determine the time period of distinct seizure stages such as ictal, pre ictal, inter ictal, and post ictal. On the transformed domain, we have utilized three different classifiers- DTC, RFC, and KNN to determine the accuracy of our detection technique in both 2 class (seizure and non-seizure) and 4 class (seizure and class (ictal, preictal, inter ictal and post ictal). Our accuracy rate in 2 class classification is extremely outstanding, with 97.22% in DTC, 100% in RFC, and 82.33% in KNN. On the other hand, our accuracy in four class classification is 91.67% in both DTC and RFC, and 80.56% in KNN. This better accuracy rate clearly demonstrates how our proposed model outperforms the state-of-the-art models.

Bibliography

[1] M. Z. Parvez and M. Paul, "Seizure prediction using undulated global and local features," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 208–217, Jan. 2017, ISSN: 0018-9294, 1558-2531. DOI: 10.1109/TBME.2016.2553131.

[2] ——, "Epileptic seizure detection by analyzing eeg signals using different transformation techniques," *Neurocomputing*, vol. 145, pp. 190–200, Dec. 2014, ISSN: 09252312. DOI: 10.1016/j.neucom.2014.05.044.

[3] E. Magiorkinis, A. Diamantis, K. Sidiropoulou, and C. Panteliadis, "Highights in the history of epilepsy: The last 200 years," *Epilepsy Research and Treatment*, vol. 2014, pp. 1–13, Aug. 2014, ISSN: 2090-1348, 2090-1356. DOI: 10.1155/2014/582039.

[4] R. A. Gross, "A brief history of epilepsy and its therapy in the western hemisphere," *Epilepsy Research*, vol. 12, no. 2, pp. 65–74, Jul. 1992, ISSN: 09201211. DOI: 10.1016/0920-1211(92)90028-R.

[5] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad, "A review of channel selection algorithms for eeg signal processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2015, no. 1, p. 66, Dec. 2015, ISSN: 1687-6180. DOI: 10.1186/s13634-015-0251-9.

[6] B. J. Baars and N. M. Gage, *Cognition, brain, and consciousness: introduction to cognitive neuroscience*, 2nd ed. Academic Press/Elsevier, 2010, ISBN: 9780123750709.

[7] R. Panda, P. S. Khobragade, P. D. Jambhule, S. N. Jengthe, P. Pal, and T. K. Gandhi, "Classification of eeg signal using wavelet transform and support vector machine for epileptic seizure diction," in *2010 International Conference on Systems in Medicine and Biology*, Dec. 2010, pp. 405–408. DOI: 10.1109/ICSMB.2010.5735413.

[8] M. Z. Parvez, M. Paul, and M. Antolovich, "Detection of pre-stage of epileptic seizure by exploiting temporal correlation of emd decomposed eeg signals," *Journal of Medical and Bioengineering*, vol. 4, no. 2, pp. 110–116, 2015, ISSN: 23013796. DOI: 10.12720/jomb.4.2.110-116.

[9] Z. Lasefr, S. S. V. N. R. Ayyalasomayajula, and K. Elleithy, "Epilepsy seizure detection using eeg signals," in *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, IEEE, Oct. 2017, pp. 162–167, ISBN: 9781538611043. DOI: 10.1109/UEMCON.2017.8249018. [Online]. Available: http://ieeexplore.ieee.org/document/8249018/.

[10] M. A. Hadj-Youcef, M. Adnane, and A. Bousbia-Salah, "Detection of epileptics during seizure free periods," in *2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA)*, May 2013, pp. 209–213. DOI: 10.1109/WoSSPA.2013.6602363.

[11] S.-F. Liang, H.-C. Wang, and W.-L. Chang, "Combination of eeg complexity and spectral analysis for epilepsy diagnosis and seizure detection," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, no. 1, p. 853 434, Dec. 2010, ISSN: 1687-6180. DOI: 10.1155/2010/853434.

[12] H. Ocak, "Optimal classification of epileptic seizures in eeg using wavelet analysis and genetic algorithm," *Signal Processing*, vol. 88, no. 7, pp. 1858–1867, Jul. 2008, ISSN: 01651684. DOI: 10.1016/j.sigpro.2008.01.026.

[13] R. B. Pachori, "Discrimination between ictal and seizure-free eeg signals using empirical mode decomposition," *Research Letters in Signal Processing*, vol. 2008, pp. 1–5, 2008, ISSN: 1687-6911, 1687-692X. DOI: 10.1155/2008/293056.

[14] V. Bajaj and R. B. Pachori, "Classification of seizure and nonseizure eeg signals using empirical mode decomposition," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 6, pp. 1135–1142, Nov. 2012, ISSN: 1089-7771, 1558-0032. DOI: 10.1109/TITB.2011.2181403.

[15] F. Duman, N. Özdemir, and E. Yildirim, "Patient specific seizure prediction algorithm using hilbert-huang transform," in *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*, Jan. 2012, pp. 705–708. DOI: 10.1109/BHI.2012.6211680.

[16] X. Zhang, W. Diao, and Z. Cheng, "Wavelet transform and singular value decomposition of eeg signal for pattern recognition of complicated hand activities," in *Digital Human Modeling*, V. G. Duffy, Ed. Springer Berlin Heidelberg,

2007, vol. 4561, pp. 294–303, ISBN: 9783540733188. DOI: 10.1007/978-3-540-73321-8_35. [Online]. Available: http://link.springer.com/10.1007/978-3-540-73321-8_35.

[17] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 9, pp. 1545–1551, Sep. 2007, ISSN: 0018-9294. DOI: 10.1109/TBME.2007.891945.

[18] S. Bezobrazova and V. Golovko, "Comparative analysis of forecasting neural networks in the application for epilepsy detection," in *2007 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, Sep. 2007, pp. 202–206. DOI: 10.1109/IDAACS.2007.4488405.

[19] K. Polat and S. Güneş, "Classification of epileptiform eeg using a hybrid system based on decision tree classifier and fast fourier transform," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 1017–1026, Apr. 2007, ISSN: 00963003. DOI: 10.1016/j.amc.2006.09.022.

[20] M. Z. Parvez and M. Paul, "Epileptic seizure prediction by exploiting spatiotemporal relationship of eeg signals using phase correlation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 158–168, Jan. 2016, ISSN: 1534-4320, 1558-0210. DOI: 10.1109/TNSRE.2015.2458982.

[21] S. R. Hirekhan, R. Manthalkar, and S. Phutke, "The detrended fluctuation analysis of eeg signals: A meditation-based study," in *Computing, Communication and Signal Processing*, B. Iyer, S. Nalbalwar, and N. P. Pathak, Eds. Springer Singapore, 2019, vol. 810, pp. 771–780, ISBN: 9789811315121. DOI: 10.1007/978-981-13-1513-8_78. [Online]. Available: http://link.springer.com/10.1007/978-981-13-1513-8_78.

[22] R. Hardstone, S.-S. Poil, G. Schiavone, R. Jansen, V. V. Nikulin, H. D. Mansvelder, and K. Linkenkaer-Hansen, "Detrended fluctuation analysis: A scale-free view on neuronal oscillations," *Frontiers in Physiology*, vol. 3, 2012, ISSN: 1664-042X. DOI: 10.3389/fphys.2012.00450. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fphys.2012.00450/abstract.

[23] Y. A. Djawad, D. Attwood, J. Kiely, and R. Luxton, "The application of detrended fluctuation analysis to assess physical characteristics of the human cell line ecv304 following toxic challenges," *Sensing and Bio-Sensing Research*, vol. 23, p. 100 269, Apr. 2019, ISSN: 22141804. DOI: 10.1016/j.sbsr.2019.100269.

[24] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. Stanley, “Multifractal detrended fluctuation analysis of nonstationary time series,” *Physica A: Statistical Mechanics and its Applications*, vol. 316, no. 1–4, pp. 87–114, Dec. 2002, ISSN: 03784371. DOI: 10.1016/S0378-4371(02)01383-3.

[25] T. N. J. Kawe, S. M. Shadli, and N. McNaughton, “Higuchi's fractal dimension, but not frontal or posterior alpha asymmetry, predicts pid-5 anxiousness more than depressivity,” *Scientific Reports*, vol. 9, no. 1, p. 19 666, Dec. 2019, ISSN: 2045-2322. DOI: 10.1038/s41598-019-56229-w.

[26] S. Kesić and S. Z. Spasić, “Application of higuchi's fractal dimension from basic to clinical neurophysiology: A review,” *Computer Methods and Programs in Biomedicine*, vol. 133, pp. 55–70, Sep. 2016, ISSN: 01692607. DOI: 10.1016/j.cmpb.2016.05.014.

[27] S.-H. Oh, Y.-R. Lee, and H.-N. Kim, “A novel eeg feature extraction method using hjorth parameter,” *International Journal of Electronics and Electrical Engineering*, pp. 106–110, 2014, ISSN: 2301380X. DOI: 10.12720/ijeee.2.2.106-110.

[28] O. Turk, M. Seker, V. Akpolat, and M. S. Ozerdem, “Classification of mental task eeg records using hjorth parameters,” in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, IEEE, May 2017, pp. 1–4, ISBN: 9781509064946. DOI: 10.1109/SIU.2017.7960608. [Online]. Available: http://ieeexplore.ieee.org/document/7960608/.

[29] F. S. Bao, X. Liu, and C. Zhang, “Pyeeg: An open source python module for eeg/meg feature extraction,” *Computational Intelligence and Neuroscience*, vol. 2011, pp. 1–7, 2011, ISSN: 1687-5265, 1687-5273. DOI: 10.1155/2011/406391.

[30] R. D. Phillips, L. T. Watson, R. H. Wynne, and C. E. Blinn, “Feature reduction using a singular value decomposition for the iterative guided spectral class rejection hybrid classifier,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, no. 1, pp. 107–116, Jan. 2009, ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2008.03.004.

[31] S. R. Ara, S. K. Bashar, F. Alam, and M. K. Hasan, “Emd-dwt based transform domain feature reduction approach for quantitative multi-class classification of breast lesions,” *Ultrasonics*, vol. 80, pp. 22–33, Sep. 2017, ISSN: 0041624X. DOI: 10.1016/j.ultras.2017.04.006.

[32] J. Vlach, P. Rajmic, J. Prinosil, J. Vyoral, and I. Mica, "Optimized discrete wavelet transform to real-time digital signal processing," in *Personal Wireless Communications*, R. Bestak, B. Simak, and E. Kozlowska, Eds. Springer US, 2007, vol. 245, pp. 514–520, ISBN: 9780387741581. DOI: 10.1007/978-0-387-74159-8_51. [Online]. Available: http://link.springer.com/10.1007/978-0-387-74159-8_51.

[33] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, no. 86, pp. 2579–2605, 2008, ISSN: 1533-7928.

[34] F. S. Nahm, "What the p values really tell us," *The Korean Journal of Pain*, vol. 30, no. 4, pp. 241–242, Oct. 2017, ISSN: 2005-9159, 2093-0569. DOI: 10.3344/kjp.2017.30.4.241.

[35] H. H. Patel and P. Prajapati, "Study and analysis of decision tree based classification algorithms," *International Journal of Computer Sciences and Engineering*, vol. 6, no. 10, pp. 74–78, Oct. 2018, ISSN: 23472693. DOI: 10.26438/ijcse/v6i10.7478.

[36] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, "Random forests and decision trees," *International Journal of Computer Science Issues(IJCSI)*, vol. 9, Sep. 2012.

[37] P. Cunningham and S. J. Delany, *K-nearest neighbour classifiers*, 2007.

[38] A. Shoeb, "Chb-mit scalp eeg database," 2010, Type: dataset. DOI: 10.13026/C2K01R. [Online]. Available: https://physionet.org/content/chbmit/.

[39] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, Jun. 2000, ISSN: 0009-7322, 1524-4539. DOI: 10.1161/01.CIR.101.23.e215. [Online]. Available: https://www.ahajournals.org/doi/10.1161/01.CIR.101.23.e215.

[40] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, 1995, ISBN: 9780195058239.

[41] A. Sharmila and P. Mahalakshmi, "Wavelet-based feature extraction for classification of epileptic seizure eeg signal," *Journal of Medical Engineering Technology*, vol. 41, no. 8, pp. 670–680, Nov. 2017, ISSN: 0309-1902, 1464-522X. DOI: 10.1080/03091902.2017.1394388.

[42] F. A. Shirazi and M. J. Mahjoob, "Application of discrete wavelet transform (dwt) in combustion failure detection of ic engines," in *2007 5th International Symposium on Image and Signal Processing and Analysis*, Sep. 2007, pp. 482–486. DOI: 10.1109/ISPA.2007.4383741.