Multivariate multiscale entropy based Schizophrenia detection from multichannel EEG signals.

Dissertation

By

Harit Yadav
17MI418

DEPARTMENT OF
ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY HAMIRPUR
HAMIRPUR (H.P.) - 177005 (INDIA)

May, 2022

राष्ट्रीय प्रौद्योगिकी संस्थान हमीरपुर
हमीरपुर (हि॰प्र॰) – 177 005 (भारत)
NATIONAL INSTITUTE OF TECHNOLOGY HAMIRPUR
HAMIRPUR (H.P.) - 177 005 (INDIA)
(An Institute of National Importance under Ministry of HRD)

DECLARATION

I hereby certify that the research presented in this dissertation titled **"Multivariate multiscale entropy based Schizophrenia detection from multichannel EEG signals"** in partial fulfilment of the requirements for the degree of **Master of Technology and submitted in the Department of Electronics and Communication Engineering**, National Institute of Technology Hamirpur, is an authentic record of my own work done between July 2021 and June 2022 under the guidance of**Dr.Abhijit Bhattacharyya, Assistant Professor**, E&CED, National Institute of Technology Hamirpur. This thesis contains material that has not been submitted for any other degree at this or any other Institute/University.

Signature of Candidate

Harit Yadav

17MI418

This is to confirm that the student's stated declaration is true to the best of my knowledge.

Signature of Supervisor

Date: 08-06-2022

Dr.Abhijit Bhattacharyya,

Assistant Professor, E&CED

The M.Tech viva-voice examination of Harit Yadav has been held on _____
.

Signature of **Supervisor**

Signature of **Internal Examiner** *Signature of* **External Examiner**

ACKNOWLEDGEMENT

During the course of completing my dissertion, I met various individuals who helped me in my work; they deserves special acknowledgement. And I'm very happy to thank each of them.

First and foremost, i'd like to thank my dissertation supervisor, Dr. Abhijit Bhattacharyya, for giving me great advice, constant support, and encouragement throughout the whole process. It's a great chance for me to learn from his experience and finish my thesis under his guidance. Working with him was such a great experience that i will never forget it.

I'm grateful to Prof. Hiralal Murlidhar Suryawanshi, Director of the NIT Hamirpur (NITH), Dr. Gargi Khanna, Head (HOD) and all faculty members of the ECE department for all their help with my thesis. I'd want to express my gratitude to all of my friends and classmates for making my experience at NIT Hamirpur so special and memorable.

I will always be thankful to my parents for all the care they have given me throughout my life and while i worked on this dissertation.

I am thankful to God for granting me the chance, confidence, and ability to finish this dissertation effectively.

Place: **Harit Yadav**

NIT Hamirpur **17MI418**

ABSTRACT

The purpose of this research is to recognise Schizophrenia from multichannel EEG signals. Schizophrenia recognition through EEG signals is a precise, cheap, short, and simple solution. This work offer a novel multivariate multiscale entropy based feature extraction algorithm for automatic Schizophrenia identification from features/entropies generated from multichannel EEG signals.

The proposed method work in four phases. In the first phase, segmentation and preprocessing are applied o collected multichannel EEG signals. In the second phase, multi-variate variational mode decomposition is used for generating multi-variate modulated oscillations inherent in multi-channel EEG signals. In the third phase, the multi-variate multiscale entropy calculation technique is used to generate features from these MMOs for sample-entropy, approximation-entropy, permutation-entropy and renyi-entropy. Finally, fourteen different machine learning (ML) classifiers is used to classify Schizophrenia paranoid and Healthy contol.

The proposed approach is used to evaluate 19-channel EEG signal recording for 14 SZ patients and 14 HC participants.The results of the study show that the proposed method outperform other Schizophrenia detection methods. We achieved highest 97.1% accuracy, 97.12% sensitivity, 97.09% Roc-Auc, 97.09% specificity, 97.1% f1-score and 97.59% precision with ensemble subspace KNN (ESK) classifier. Our study reveal that the multivariate multiscale entropy based feature extraction for mulcichannel EEG signals achieve good performance for Schizophrenia classification.

Contents

List of Figures

List of Tables

List of Abbreviations

ApEn	Approximation Entropy
AUC	Area Under Curve
C2F	Coarse-to-Fine
CKNN	Cubic K-Nearest Neighbours
CNN	Convolutional neural network
CSVM	Cubic Support Vector Machine
CWT	Continuous Wavelet Transform
DT	Decision Tree
EBAT	Ensemble Bagged-Tree
EBT	Ensemble Boosted-Tree
EEG	Electroencephalogram
EMD	Empirical-Mode Decomposition
ERT	Ensemble Rusboosted Tree
ESD	Ensemble Subspace Discriminant Analysis
ESK	Ensemble Subspace K-Nearest Neighbour
EWT	Empirical Wavelet Transform
F2C	Fine-to-Coarse
FFT	Fast Fourier Transform
FKNN	Fine K-Nearest Neighbours
FNR	False Negative Rate
FPR	False Positive Rate
ICD	International Classification of Disease
IMF	Intrinsic Mode Functions
KNN	K-Nearest Neighbours
LDA	Linear Discriminant Analysis

LR	Linear Regression
LSTM	Long Short-Term Memory
LSVM	Linear Support Vector Machine
ML	Machine Learning
MME	Multivariate Multiscale Entropy
MMO	Multivariate Modulated Oscillations
MVMD	Multivariate Variational Mode Decomposition
PeEn	Permutation Entropy
PSD	Power Spectral Density
QDA	Quadratic Discriminant Analysis
QSVM	Quardic Support Vector Machine
RBF	Radial Basis Function
ReEn	Renyi Entropy
RF	Random Forest
ROC	Receiver Operating Curve
SaEn	Sample Entropy
SPWVD	Smoothed Pseduo-winger-ville Distribution
STFT	Short Time Fourier Transform
SVM	State Vector Machine
TF	Time Frequency
VMD	Variational Mode Decomposition
WKNN	Weighted K-Nearest Neighbours

Chapter 1

INTRODUCTION

1.1 Introduction

Schizophrenia is one of the most serious mental illnesses, and it impacts a person's ability to think, feel, behave, express emotions, understand reality, and socialize with others. [1] Schizophrenia affects about 24 million (around 0.32% of the world population) people worldwide. The term "schizophrenia" means "fragmented mind". According to research, schizophrenia patients experience a range of symptoms [2], like disorganised thoughts, hallucinations, and other psychological disorders. According to scientific analysis, schizophrenia is produced by a combination of various genetic and environmental aspects. Psychological and social variables may play a part in the development and advancement of schizophrenia. According to the World Health Organization, heavy cannabis use also raises the risk of schizophrenia. Typically, schizophrenia occurs in late adolescence or early infancy (less commonly affects children or after the age of 40 years). The peak age group for men is 20 to 38 years, while for women it is 26 to 32 years. Men are diagnosed with schizophrenia at a ratio of 1.40:1 more frequently than women.

Although there is no permanent treatment for schizophrenia, The goal of schizophrenia treatment is to reduce the symptoms and reduce symptoms returning. [3] Treatment for schizophrenia includes Primary medication called "antipsychotic" or first-generation (Chlorpromazine, Fluphenazine, Haloperidol, etc.) which help to relieve most symptoms, "atypical" or second-generation medicines (Aripiprazole, Brexpiprazole, Cariprazine) which used to treat schizophrenia, Coordinated Speciality

Care (CSC), Psychosocial therapy and Electro-Convulsive Therapy (ECT).

Accurate prediction of schizophrenia can help to reduce the suffering of patients by starting early treatment [4] and also lower the risk of drug abuse. Research shows that people with SZ die 2-3 times younger than the regular population. Which are caused by some physical illnesses [5] like metabolic, cardiovascular and bacterial diseases which are curable with additional care.

There are several methods for diagnosing schizophrenia, including interviews, Magnetic Resonance Imaging (MRI) [6–8], Electroencephalography (EEG) [9–12], Computed Tomography (CT) [13], and so on. The interview diagnosis process is one of them because it is a manual process that is time-consuming, burdensome, error-prone, and biased. Compared to EEG approaches, neuroimaging techniques (e.g. MRI, CT) are expensive and require high-level imaging techniques, hefty machines, and additional recording and calculation time. Currently, EEG has emerged as the standard for schizophrenia diagnosis due to its non-invasiveness, affordable and high temporal resolution in comparison to other tests [14]. The EEG technique measures tiny electrical signals generated by a large number of neurons in the head by implanting electrodes to standard position of the scalp [15]. EEG recordings contain a large volume of dynamic data for studying brain activities. There is currently no automatic, quick, and precise method to detect schizophrenia from EEG data.

1.2 EEG usage in Schizophrenia

EEG signal imaging is mostly used for the detection of neurological disorders, and it is also commonly used for the diagnosis of schizophrenia disorder. According to [16], the response time of schizophrenia patients was higher than healthy control patients. One of the main reasons EEG is used to diagnose schizophrenia is that it is easy to test and can be done in almost any psychiatric setting because almost every patient can handle it.

In [17], a status report of EEG abnormalities is submitted to diagnose schizophrenia in patients. For that, they studied the evolution of spectral EEG anomalies. The meta-analysis was constrained to studies that examined spectral power between a

group of schizophrenia patients and a healthy control group. Data collection required two distinct groups (one with schizophrenia and the other with healthy controls) to identify distinctive characteristics. Schizophrenia patients were supposed to have a higher delta, higher theta, lower-alpha, and higher beta powers. After much research, it was found that people with schizophrenia have much higher levels of activity in the lower spectrum (slow waves). According to many researches, frontal regions of the brain are more affected by slow-wave abnormalities (especially delta increase). According to [8] conclusions, the strong biomarker of schizophrenia is delta excess (and to a smaller degree theta excess). When diagnosing schizophrenia, electrodes should be placed in the frontal, parietal, central, temporal, and occipital regions of the brain.

Event-related EEG reveals that neural rythms and their synchronisation are important processes for interneural transmission and the binding of data processed by distributed brain areas [18]. Several EEG studies aim to acquire new insights into the neuropathology behind brain abnormalities in neuropsychiatric disorders, hence suggesting the use of EEG signal data analysis in the study of schizophrenia recognition.

1.3 Inspiration

Schizophrenia is a chronic psychiatric disorder that has been proven to have severe negative effects on both individuals and society as a whole. The disease is associated with several negative effects, such as decreased quality of life, depression, decreased physical health, stigma, a higher suicide risk, homelessness, and drug misuse. The disease also has societal effects, such as increased health care expenditure and care burden. In addition, schizophrenia imposes a relatively massive burden on the world economy due to the severity of the disorder and the lack of adequate treatment alternatives. Therefore, a good understanding of the symptoms and functional results caused by schizophrenia disorder is required for the development of more effective treatments and the improvement of outcomes for those with the disease. Our current understanding of schizophrenia disorder is particularly insufficient in regards to the disorder's negative symptoms, such as delusions and hallucinations, which have

received great attention in many researches and are often treated with pharmacologic treatments. But negative symptoms, like apathy, anhedonia, alogia, and social withdrawal, cannot be adequately addressed by any method. Negative symptoms are very burdensome for patients since they are often chronic, continue during remission, and are connected to overall poor activity and quality of life. Since there is no permanent cure, early detection can help to minimise the disease's impact. Therefore, research on the detection of schizophrenia is required. There are various methods for diagnosing schizophrenia, like magnetic resonance imaging (MRI), electroencephalography (EEG), and computed tomography (CT), among others. There is no reliable, accurate, real-time approach for recognising schizophrenia.

In this study, we like to create a Schizophrenia detection model with high sensitivity, robustness, and less required time. We aim to study EEG signals, as they have high data and cheap relative to other diagnostic technologies. To produce features, feature generation algorithms will be applied, and machine learning will be utilised to train and classify them.

1.4 Aim

The following are the main objectives of this research work:

- To create an EEG-signal based Schizophrenia detection algorithm that is effective and precise.
- To study and implement multivariate variation mode signal decomposition on EEG data.
- To Calculate different types of entropies with a multivariate multiscale technique on decomposed EEG-signals to generate features.
- To train different type of machine learning classifiers on these generated features.
- Analyse the classifications of schizophrenia obtained by several machine learning classifiers.

1.5 Contribution

These are the major contributions for the suggested model:

- Investigate several EEG electrodes in order to enable quick processing, down-sampling, and identification of brain areas impacted by schizophrenia.
- MVMD is used to study the multichannel EEG signal decomposition approach for extracting MMOs.
- Generating features by calculating different types of entropies with multivariate multiscale method on MMOs generated by MVMD.
- Various machine learning classifiers are used for training and classification of these generated features.
- Comparing the performance of these trained ML classifiers for Schizophrenia recognition.

1.6 Work Outlines

- The **second chapter** reviews current research on the EEG signal-based schizophrenia recognition model. It also investigates various EEG signal decomposition strategies. This chapter provides an overview of numerous approaches for extracting features.
- The **third chapter** explains all methodological elements of the proposed technique. This section discusses the EEG datasets utilised by the proposed approach. This chapter discusses MVMD decomposition method and multivariate multiscale entropy analysis of EEG signals in details. This chapter also describes the various entropies calculated in this study.
- The **fourth chapter** explains the experimental design process in this proposed technique. In addition, it explains the performance measures used to examine the model.

- The **fifth chapter** analyzes the results of the suggested method using publicaly available datasets. In this chapter, the outcomes of all techniques are discussed.

- The **sixth chapter** provides a summary of the main findings that are derived from this work.

Chapter 2

RELATED WORK

2.1 Related Work

In recent years, various studies have been conducted to obtain important features from EEG signals to build a Schizophrenia detection model. das et al. [19] developed a new model for Schizophrenia recognition using multi-variate iterative filtering as feature extraction, t-test as feature selection and for identifying features machine learning classifiers such as KNN, LDA, SVM, DT are used. And they achieved the highest 98.9% accuracy with the SVM(cubic) classifier in this method. Buettner et al. [20] reported a feature extraction process where EEG signals were separated into five frequency bands using FFT. They fed these obtained features as input to a random forest (RF) classification algorithm to differentiate schizophrenia and healthy control classes. They attained the highest 96.77% accuracy in this method. In literature [21], Khare et al. developed a method that combines time-frequency analysis and a convolutional neural network(CNN) classifier. EEG signals are processed with CWT, STFT, and SPWVD to generate TF images. These 2d images are given to pretrained transfer learning models (AlexNet, ResNet50, CNN and VGG16). They attained the highest 93.36% accuracy using SPWVD based TF plots and the CNN model. Shoeibi et al. [22], divide EEG signals into 25 second time frames and normalized them using the z-score test. The signals are trained using convolutional deep learning (1d-CNN) and then CNN- LSTM. This method achieves 99.25% accuracy with CNN-LSTM. Siuly et al. [23], in this study EWT is used for decomposing EEG signals into small IMFs. from these IMFs, twenty-two statistical features are

calculated and five features were selected using Kruskal- Wallis test and trained using ML classifiers. This method achieves the highest 93.21% accuracy with an ensemble bagged tree classifier. Aslan et al. [24], convert the raw EEG data into 2D TF plots by using STFT and feeding these images to a pre-trained VGG-16 model. This experiment achieves 95% accuracy. Vicnesh et al. [25], This approach used an eleven layers CNN networks to extract features from EEG signals and classify them. This model has an accuracy of 98.07 % for non-subject assessment and 81.26% for subject-based assessment. According to Shalbaf et al. [26], CWT is used to convert EEG signals into TF images. These TF images are fed to pre-trained transfer learning models: inception-v3, VGG-19 and Resnet- 18. This method achieves the highest accuracy of 98.60% with Resnet18-SVM. According to Shalbaf et al. [27], this approach applies a combination of independent component analysis and a random forest classifier on EEG signal data. This method achieves an accuracy of 96.77%. Sabeti et al. [28], used a combination of linear discriminant analysis (LDA) and adaptive boosting classifiers on collected EEG data. This method achieves an accuracy of 86 to 90%, which is further improved to 91 to 92% through computational adjustments. Krishanan et al. [29] uses multi-variate empirical mode decomposition to convert EEG signal into IMFs, and extract five different types of entropies from IMFs. They achieved maximum accuracy of 93% using the RSVM classifier.

2.2 Overview

A deep learning-based classification usually require more system resources and more time to develop than ML classifiers. Other feature extraction methods use the univariate decomposition method to decompose multi-channel EEG data individually. As a result, mutual information between different channels will be lost. Handmade methods for extracting features from EEG data take a lot of time and work to choose the best methods for each set of EEG data. MRI based approaches also provide better results, but image techniques are costly and time-consuming. We are using EEG signal data since it's cheap and can be recorded with minimal instruments.

Chapter 3

SUGGESTED METHOD

3.1 Suggested Method

We suggested a robust multivariate multiscale entropy calculation based feature extraction technique for Schizophrenia recognition by applying the machine learning-based classification algorithms. The benefit of using multivariate multiscale entropy is that it also calculates mutual information between multichannel and multimode, which is lost when using univariate methods. The multivariate multiscale method can be easily tuned as per requirements. Machine learning-based classification algorithms can be trained faster using fewer computer resources, and get good results even on small datasets in comparison to deep learning algorithms. EEG data are analysed using several entropies (sample entropy, approximation entropy, permutation entropy, and renyi entropy) to evaluate schizophrenia patterns. Considering that different channels contain mutual information or produce Schizophrenia patterns, we have researched entropies on all scales to develop a Schizophrenia detection model with minimal computation costs and work in real-time.

The proposed method work in four phases: segmentation and pre-process of collected EEG data, MMO's generation from multichannel EEG signals, feature extraction from MMOs, and classification of Schizophrenia. For generating MMOs innate in multichannel EEG signals, we use the MVMD technique given in [30]. MVMD uses a model for multi-variate modulated oscillations based on a common frequency component to extract band-limited modes comprising data-intrinsic modulated oscillations. We use the multivariate multiscale entropy calculation technique

to extract features from these MMOs. The obtained features are fed to machine learning classifiers for Schizophrenia classification.

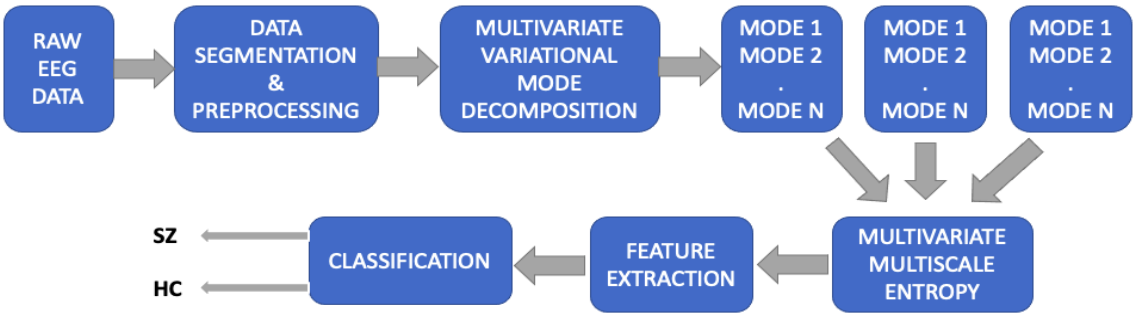


Figure 3.1: Schematic diagram of the proposed approach for recognising schizophrenia from EEG signals.

3.2 EEG Dataset & Preprocessing

EEG data is the recording of the electrical activity of the brain using scalp electrodes (non-invasive). In this work, we used a publicly available Schizophrenia dataset [31] from the Institute of Psychiatry and Neurology in Warsaw, Poland [32]. This dataset contains multichannel EEG signals recording of 28 patients. In which 14 patients including seven males(age: 27.9 $\pm$ 3.3 years) and seven females(age: 28.3 $\pm$ 4.1 years) with schizophrenia disorder, who were hospitalized at the Institute of Psychiatry and Neurology in Warsaw, Poland, and 14 healthy controls including seven males(age: 26.8 $\pm$ 2.9) and seven female(ages: 28.7 $\pm$ 3.4 years). Patients met ICD-10 Schizophrenia disorder criteria (category F20.0). EEG data were recorded for fifteen minutes for all subjects in the "eyes-closed resting state" condition. Data were sampled at 250 Hz using the typical 10–20 EEG montage with 19 EEG channels(Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2) for 15 minutes. Where Fcz is used as a reference electrode for 19 channels. 30 second EEG segments were used to analyse without any artefacts (eye movements, muscle contractions, and cardiac activity) EEG signals. This was followed by second-order Butterworth filters in the following physiological frequency ranges: alpha(8 to 12.5Hz), beta(13 to 30Hz), delta(2 to 4 Hz), theta(4.5 to 7.5 Hz) and gamma(30

Table 3.1: Detailed information about Schizophrenia database.

No. of subjects	**Category**	**Gender Distribution**	**Average Age**	**Sampling rate(Hz)**	**Available EEG channels**
14	HC	7M 7F	26.8 28.7	250	Fp1 Fp2 F7 F3 Fz F4 F8 T3 C3 Cz C4 T4 T5 P3 Pz P4 T6 O1 O2
14	SZ	7M 7F	27.9 28.3	250	Fp1 Fp2 F7 F3 Fz F4 F8 T3 C3 Cz Cz T4 T5 P3 Pz P4 T6 O1 O2

to 45 Hz) for each individual channels. This dataset excludes patients with the following conditions: pregnancy, organic brain pathology, severe neurological illnesses (such as epilepsy, Alzheimer's, or Parkinson's), the existence of a general medical condition, and extremely early stage Schizophrenia, i.e. first episode of schizophrenia. The control group's gender and age were matched to those of the 14 patients who completed the research. This dataset included only for subjects above 18 years and who have completed 7 days meditation before diagnosis.

EEG signals are divided into 25-second blocks, without overlapping. Segmentation gives 1142 EEG patterns comprised of 6250 x 19 sample points. There are 516 EEG patterns in the Healthy Control group, and there are 626 EEG patterns in the Schizophrenia group. The notch filter is used for power line index removing. To perform this study, we are using all 19 channels available in the dataset. After segmentation, the patient's raw EEG data is spitted into small but equal segments. We will be checking which brain regions or electrodes are more affected by Schizophrenia. Fig.3.2 and Fig.3.3 shows the Sample recording of EEG signal representation and Power spectral density for T4(temporal), F4(frontal) and C4(central) electrodes in the Schizophrenia paranoid and Healthy control respectively. The figure show that visual analysis cannot distinguish EEG waveforms significantly. To provide appropriate data for Schizophrenia detection, we need powerful signal processing and feature extraction approaches.

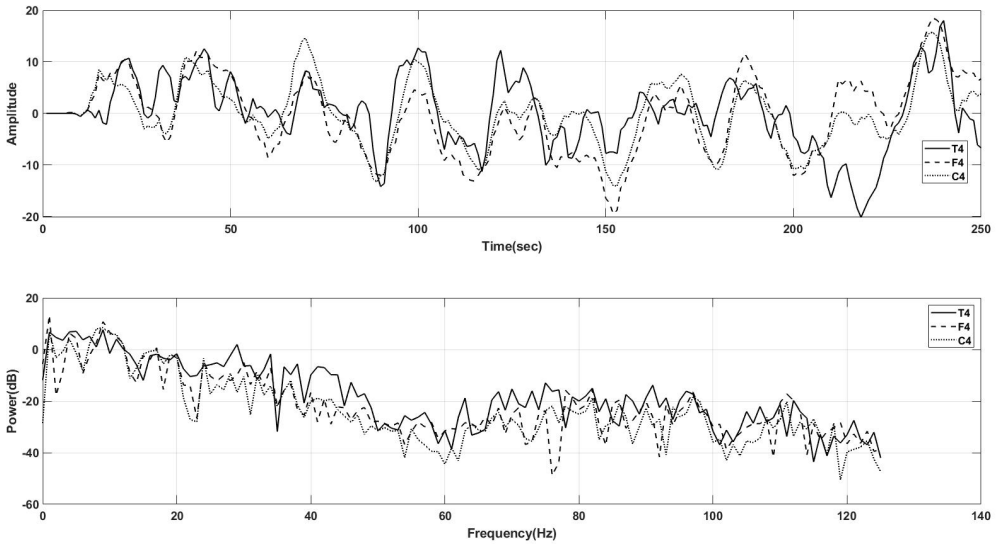


Figure 3.2: Visualization of a sample from EEG signals (top) and Power spectral density (bottom) for T4(temporal), F4(frontal) and C4(central) electrodes in the Schizophrenia paranoid.

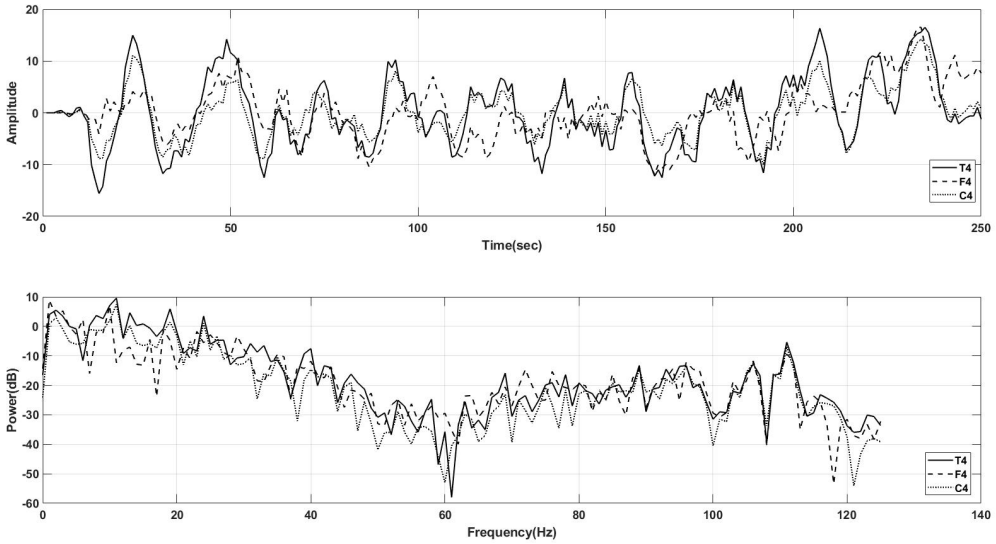


Figure 3.3: Visualization of a sample from EEG signals (top) and Power spectral density (bottom) for T4(temporal), F4(frontal) and C4(central) electrodes in the Healthy Control.

3.3 Multi-variate Variational Mode Decomposition

Multi-variate Variational Mode Decomposition (MVMD) is a generalisation of Variational Mode Decomposition (VMD) [33], that is proposed to build a data-driven

and mathematically solid framework for signal decomposition and time-frequency assessment of multi-channel non-stationary data. For a multivariate input signal $v(t) = [v_1(t), v_2(t), v_3(t), ..., v_M(t)]$ consisting of M data channels and primary goal of this technique is to extract a predetermined N MMOs (multi-variate modulated oscillations) $o_n(t)$.

$$v(t) = \sum_{n=1}^{N} o_n(t) \tag{3.1}$$

where $o_n(t)$ is defined as $o_n(t) = [o_1(t), o_2(t), ...o_N(t)]$.

The extracted MMOs $\{o_n(t)\}_{n=1}^{N}$ ensemble should meet the two conditions listed below. First, extracted modes' total bandwidth must be as little as feasible. Second, combination of all extracted modes should recover the original data precisely. For that, the related analytic signal of $r_n(t)$ is first calculated using the Hilbert transform, and its vector analytical representation is indicated by $r_+^n(t)$. Then, MMO bandwidth is calculated by moving the unilateral frequency band of all individual channels of the vector $r_+n(t)$ by harmonic mixing with a complex exponential of a common universal frequency component ϕ_n and computing the Frobenius norm^2 of the resulting matrix. As a result, a multivariate expansion of the cost function used in the VMD optimizing [33] is generated. The cost parameter L of MVMD is represented as follows:

$$L = \sum_{n} \sum_{m} \left\| \delta_t \left[o_+^{n,m}(t) e^{-j\psi t} \right] \right\|_2^2 \tag{3.2}$$

In this case, $o_+^{n,m}(t)$ is the complicated analytical modulated signal associated with the $p^(th)$ oscillation mode of m^{th} channel. The following is the solution to the MVMD parameter optimization problem:

$$\begin{aligned} &\min_{\{o_{n,m}\}\{\psi_n\}} \left\{ \sum_{n} \sum_{n} \left\| \delta_t \left[o_+^{n,m}(t) e^{-j\psi t} \right] \right\|_2^2 \right\} \\ &s.t \sum_{n=1}^{N} o_{,m}(t) = v_m(t), m = 1, 2, 3, 4..., M. \end{aligned} \tag{3.3}$$

The desired outcome for the parameter optimization issue is obtained by converting it in a variational problem with no restrictions. Using σ as a quadratic penalty term to improve the accuracy of the reconstruction and β as a Lagrangian

multiplier to make the constraints strict helps to make the optimization problem unconstrained. The augmented Lagrangian function that comes out of this is given by:

$$\begin{aligned}\mathcal{L}\left(\{o_{n,m}\},\{\psi_n\},\beta_m\right) &= \sigma\sum_n\sum_m \left\|\partial_t\left[o_+^{n,m}(t)e^{-j\psi_n t}\right]\right\|_2^2 \\ &+ \sum_L \left\| v_m(t) - \sum_n o_{n,m}(t)\right\|_2^2 \\ &+ \sum_m \left\langle \beta_m(t), v_m(t) - \sum_n o_{n,m}(t)\right\rangle .\end{aligned} \tag{3.4}$$

The ADMM solves the parameter optimization issue, which reduces this complicated issue to a series of repetitive sub-optimization problems. The following are the steps in the MVMD process:

Step 1: Intiate $\{\psi_n^1\}, \{\hat{o}_{n,m}^1\}, \hat{\beta}_m^1$, and j.

Step 2: Update the central frequencies $\{\psi_n\}$, components $\{\hat{o}_n\}$, and the Lagrange multiplier $\hat{\beta}_m$ as per requirement of equations **??**.

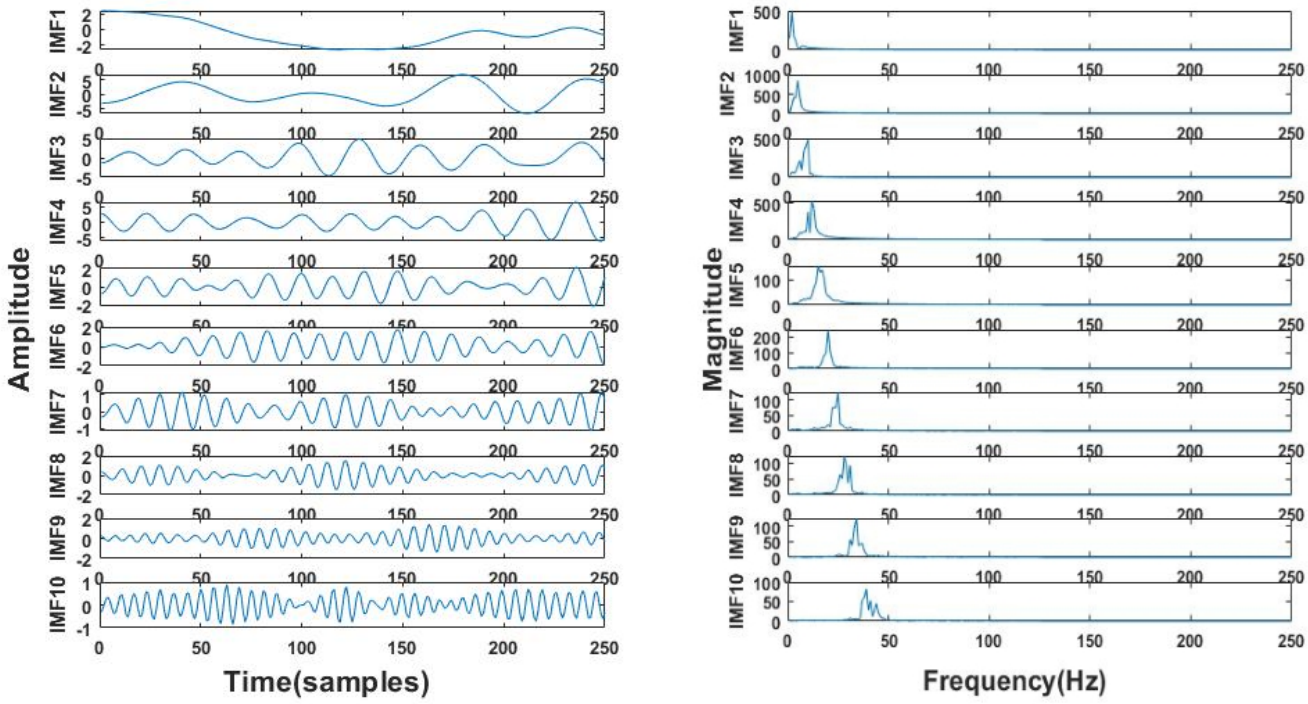
Step 3: Quit when convergence conditions are attained $\sum_k \sum_L \frac{\left\|\hat{o}_{n,m}^{j+1}-\hat{o}_{n,m}^{j}\right\|_2^2}{\left\|\hat{o}_{n,m}^{p}\right\|_2^2} < \epsilon$, otherwise go to step 2.

$$\hat{o}_{n,m}^{j+1}(\psi) = \frac{\hat{v}_m(\psi) - \sum_{i<n}\hat{o}_{i,n}^{j+1}(\psi) - \sum_{i>n}\hat{o}_{i,m}^{j}(\psi) + \frac{\beta_m(\psi)}{2}}{1+2\sigma\left(\psi-\psi_n^j\right)^2} \tag{3.5}$$

$$\psi_n^{j+1} = \frac{\sum_m \int_0^\infty \psi\left|\hat{o}_{n,m}^{j+1}(\psi)\right|^2 d\psi}{\sum_m \int_0^\infty \left|\hat{o}_{n,m}^{j+1}(\psi)\right|^2 d\psi} \tag{3.6}$$

$$\hat{\beta}_m^{j+1}(\psi) = \hat{\beta}_m^j(\psi) + \tau\left(\hat{v}_m(\psi) - \sum_n \hat{o}_{n,m}^{j+1}(\psi)\right) \tag{3.7}$$

The algorithm is ended when each subsequent set of components remains constant. Due to it's noise resilience, quasi-orthogonal rule, separation of MMOs contained in the data and mode-alignment feature, the approach is preferred. Fig.3.4 demonstrates the MMOs retrieved from the Frontal Brain area using this approach and many EEG channels.



(a) (b)

Figure 3.4: (a) Graph of IMFs computed using the MVMD technique from multi-channel EEG recordings corresponding to the Frontal Region channel for paranoid Schizophrenia and (b) their related spectrum.

3.4 Entropy Measurements

This chapter discusses the entropy measurements [34] utilised in this study's analysis. Using multivariate, multiscale entropy, the entropy measures can detect Schizophrenia patterns. In addition, we specified hyper-parameters for each entropy in order to compare results.

3.4.1 Renyi Entropy

The entropy is named after "Alfréd Rényi," who sought the broadest definition of information measures that supports the additivity of independent events.Rényi entropy is the basis of generalised dimensions in fractal dimension estimation. The definition of the Rényi entropy of order α, where $\alpha \neq 1$ and $\alpha \geq 0$, is:

$$ReEn = \frac{1}{1-\alpha} \log(\sum_{i=1}^{N} (P(x_i)^{\alpha})) \tag{3.8}$$

Here $P(x_i)$ is defined as the probability of $P(x = x_i)$. As α gets closer to zero, the [35] Rényi entropy treats all events with a probability greater than zero more equally, without regrading their probabilities. In the limit where $\alpha = 0$, the Rényi entropy becomes a logarithm of the size of X's support. As α gets closer to infinity, the highest probability events have a bigger impact on the Rényi entropy. For calculating renyi entropy, we used MATLAB implementation available at [36]. For renyi entropy calculation we used second order entropy and 6250 no. of divisions as hyperparameters.

3.4.2 Approximation Entropy

Pincus came up with ApEn in [37], and it check how predictable or regular are time series. It was made as an approximation of the Kolmogorov entropy of a process underneath it.

ApEn is defined by an algorithm that looks for repeating patterns of length m that start at sample i and have distances that are different by up to an error threshold r.

The following equation defines ApEn:

$$ApEn = \psi^m(r) - \psi^{m+1}(r) \tag{3.9}$$

Where:

$$\psi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \log(G_i^m(r)) \tag{3.10}$$

Here $G_i^m(r)$ represent number of vectors ($x_i \, \epsilon \, \Re^m$ so that $x_i = x_1, x_2, x_3, ...x_{i+m-1}$ and distance will be $d(x_i, x_j) \; < \; r$.

For calculating approximation entropy, we used MATLAB implementation available at [37, 38]. For approximation entropy calculation we used dimension = 2,and tolerance = 0.001 as hyperparameters.

3.4.3 Permutation Entropy

In their seminal paper [39], Bandt and Pompe first introduced PeEn. By comparing neighbouring values in a time series, the idea is to define an entropy measure that determines the time causality of the time-series data. PeEn is simply, Shannon

Entropy of an ordinal patterns sequence. Ordinal patterns are discrete symbols that show how entries in a time series are related to each other in terms of value and position.

Permutation entropy of order D is determined by:

$$PeEn = -\sum_{i=1}^{D!} p_i \log(p_i) \tag{3.11}$$

Where p indicates the relative frequency of every pattern of symbol sequences, called "permutation". The permutation is related to a set of m (embedding dimension) values from the original series. A time delay τ may be utilised in the creation of permutations from the original series. However, for the sake of simplicity, τ is fixed to 1 in this research, while m is a hyper-parameter to be tuned. Additional information on the PeEn's definition is available in [39]. For calculating permutation entropy, we used MATLAB implementation available at [40]. For permutation entropy calculation we used dimension = 2 as hyper-parameters.

3.4.4 Sample Entropy

SaEn was defined as a development of ApEn with the goal of addressing ApEn's bias due to counting self-matches, but it demonstrated to better statistical qualities than ApEn in many circumstances [41]. It is calculated similarly to ApEn as explained in Section 3.4.2, except the last step of computing SaEn is as follows:

$$SaEn = -\log\left(\frac{A(r)}{B(r)}\right) \tag{3.12}$$

Where B(r) is specified as the average of the number of vectors x_i^m whose distance from $d(x_i, x_j < r$ is divided by (N-m + 1).where $x_i \epsilon\ \Re^{m+1}$ defines the value of A(r) same as B(r). For calculating sample entropy, we used MATLAB implementation available at [41, 42]. For sample entropy calculation we used dimension = 2 and tolerance = 0.2 as hyper-parameters.

3.5 Multivariate Multiscale Entropy

This method was built to assess the complexities of time series at various scales. The scales are calculated using the "coarse-grained" method, in which the duration of the

"coarse-grained" data series is equal to the duration of the original data series divide by the n [43]. In reality, the technique is a linear flattening and compression of the original data series, which eliminates high-frequency elements. MMSE is attained by calculating various entropies (specified in Section 3.4) over different scales.

Our proposed MME method divided into two parts: 1) Execute the MVMD to divide signal data into affiliated IMFs at various scales, and 2) calculate the various entropies over the selectable scales. These scales are determined by consecutively reducing either the large-bands or small-bands IMFs from the source data. The method of scale selection generates two algorithms, the fine-to-coarse MME and the coarse-to-fine MME [44], which are described below. The fine-to-coarse and coarse-to-fine scales effectively represent the original signal's multiscale low-pass and high-pass filtering, respectively.

Algorithm-1: The fine-to-coarse MME

- Execute the MVMD to generate N-affiliated IMFs for all data (the residual will be last IMF).
- Select the multivariate data scales by consecutively deleting the high-frequency IMFs in order, beginning with the first IMF $S^k_{f2c}(k, m, r) = \sum_{i=k}^{N} IMF_i$, where $k \leq N$
- The MME is computed by implementing the SampEn, ApEn, ReEn, and PeEn (described in section 3.4) to each scale to get an entropy dependent on scales.These are equations to calculate MME with various entropys.
  - For Sample Entropy: $\mathbf{MME}_{f2c}(k, m, r) = SaEn(S^k_{f2c}, m, r)$
  - For Approximation Entropy: $\mathbf{MME}_{f2c}(k, m, r) = ApEn(S^k_{f2c}, m, r)$
  - For Permutation Entropy: $\mathbf{MME}_{f2c}(k, m, r) = PeEn(S^k_{f2c}, m, r)$
  - For Renyi Entropy: $\mathbf{MME}_{f2c}(k, m, r) = ReEn(S^k_{f2c}, m, r)$

Algorithm-2: The coarse-to-fine MME

- Execute the MVMD to generate N-affiliated IMFs for all data (the residual will be last IMF).

- Select the multivariate data scales by consecutively deleting the low-frequency IMFs in order, beginning with the last IMF $S^k_{c2f}(k,m,r) = \sum_{i=1}^{N-k+1} IMF_i$, where $k \leq N$

- The MME is computed by implementing the SampEn, ApEn, ReEn, and PeEn to each scale to get an entropy dependent on scales.These are equations to calculate MME with various entropys.

  - For Sample Entropy: $\mathbf{MME}_{c2f}(k,m,r) = SaEn(S^k_{c2f},m,r)$
  - For Approximation Entropy: $\mathbf{MME}_{c2f}(k,m,r) = ApEn(S^k_{c2f},m,r)$
  - For Permutation Entropy: $\mathbf{MME}_{c2f}(k,m,r) = PeEn(S^k_{c2f},m,r)$
  - For Renyi Entropy: $\mathbf{MME}_{c2f}(k,m,r) = ReEn(S^k_{c2f},m,r)$

MME generate 1142 x 190 size feature matrix for each entropy using both algorithms. After calculating features, we obtained total eight feature matrix which will act as dataset for classification. We obtained 4 feature matrix with coarse-to-fine algorithm and as well as 4 from fine-to-coarse algorithm. Further these feature matrix will be fed to machine learning classifiers.

3.6 Schizophrenia Classification

In this section, we apply classification on multi-variate multiscale entropy-based generated features for Schizophrenia and Healthy control category. For binary classification, we labelled the Schizophrenia class as 1 and the Healthy control class as 0. For this evaluation we used fourteen different machine learning binary classification algorithms: Logistic Regression (LR), Linear discriminant analysis (LDA), Quadratic discriminant analysis (QDA), Linear-support vector machine(LSVM), Quardic-support vector machine (QSVM), Cubic-support vector machine (CSVM),Fine K-nearest neighbours (FKNN), CUBIC K-nearest neighbours (CKNN), Weighted K-nearest neighbours (WKNN),Ensemble boosted tree (EBT), Ensemle bagged tree (EBAT),Ensemble subspace discriminant(ESD), Ensemble subspace KNN (ESK) and Ensemble rusboosted tree (ERT). A short explanation of these classifiers can be found in [45–52]. Table 3.2 shows parameters of different ML classifiers. We used

Baysen optimization as optimizer, Expected improvement per second as acquistition function and train each classifier over 30 iterations.

Table 3.2: Hyper-parameters used by different classifier.

Classifiers	**Parameters**	**Values**
LSVM	Kernel scale	auto
	Kernel type	Linear
	Multiclass method	1v1
	Box constraint level	1
QSVM	Kernel scale	auto
	Kernel type	Quardic
	Multiclass method	1v1
	Box constraint level	1
CSVM	Kernel scale	auto
	Kernel type	Cubic
	Multiclass method	1v1
	Box constraint level	1
FKNN	Distance metric	Euclidean
	No. of neighbours	1
	Distance weight	Equal
CKNN	Distance metric	Minkowski
	No. of neighbours	10
	Distance weight	Equal
WKNN	Distance metric	Euclidean
	No. of neighbours	10
	Distance weight	Square inverse
EBT	Learner type	Decision Tree
	Ensemble method	Adaboost
	No. of learners	30
	Learning rate	0.1
	Max. no. of splits	20
EBAT	Learner type	Decision Tree
	Ensemble method	Bagging
	No. of learners	30
	Learning rate	0.1
	Max. no. of splits	1141
ESD	Learner type	Discriminant
	Ensemble method	Subspace
	Subspace dimension	95
	No. of learners	30
ESK	Learner type	Nearest neighbour
	Ensemble method	Subspace
	Subspace dimension	95
	No. of learners	30
ERT	Learner type	Decision Tree
	Ensemble method	Rusboost
	No. of learners	30
	Learning rate	0.1
	Max. no. of splits	20

Chapter 4

RESEARCH DESIGN

4.1 Research Design

In this research, all studies are conducted using MATLAB (2021b) on a PC equipped with an eight-gen Intel i5 processor and 8GB of RAM. The computer is equipped with a 4GB NVIDIA GeForce 1050 graphics processing unit. This study used the suggested approach on multichannel EEG data. Working memory-related areas of the medial temporal and prefrontal lobes are most affected by schizophrenia. According to figure 4.1, it shows electrodes (includes A1,A2 electrodes which are not recorded in dataset) placement over different part of brain. it used EEG 10-20 electrode placement. The numbers "10" and "20" indicate that the actual distances between adjacent electrodes are 10% or 20% of the entire front–back or right–left distance of the skull, respectively. For instance, from the nasion to the inion, a measurement is made across the top of the head.

In this study, we used all 19 channels available in the dataset. EEG signals are divided into 25-second blocks, without overlapping. Segmentation gives 1142 EEG patterns comprised of 6250 x 19 sample points. There are 516 EEG patterns in the Healthy Control group, and there are 626 EEG patterns in the Schizophrenia group. We generated MMOs inherent in multichannel EEG signals, with the MVMD signal decomposition technique given by Rehman et al [30]. We used the multivariate multiscale entropy calculation technique to extract features from these MMOs for both Schizophrenia and healthy control class. The obtained features are fed to machine learning classifiers for Schizophrenia classification. To evaluate the perfor-

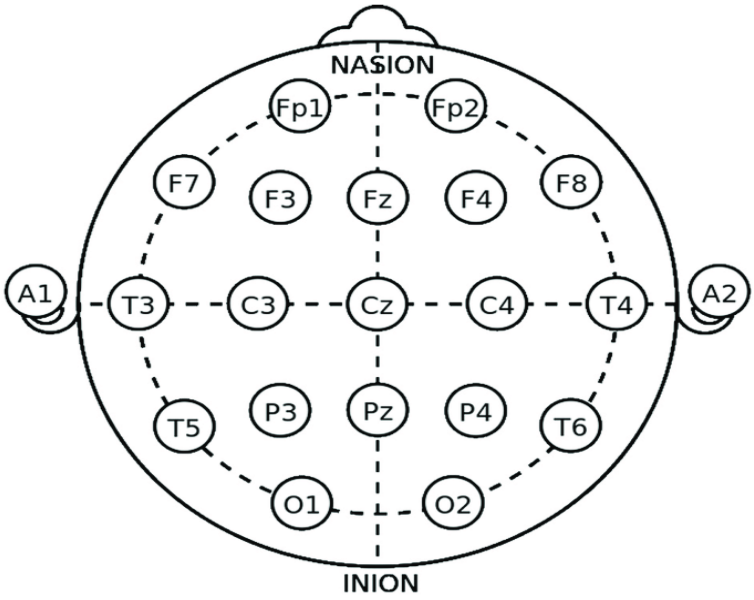
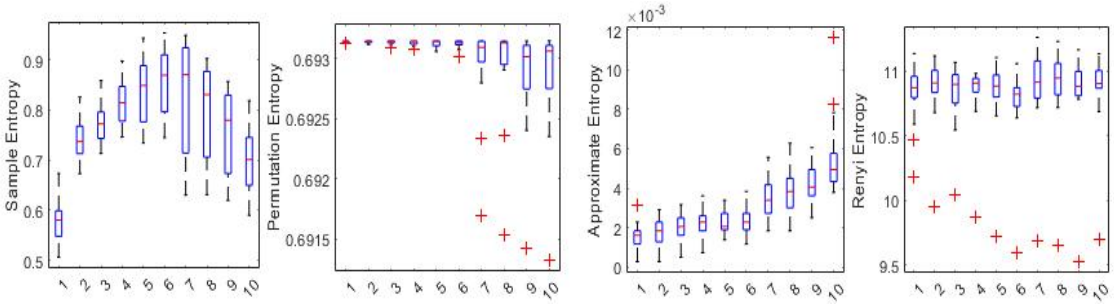


Figure 4.1: EEG electrodes placement layout according to different part of brain.

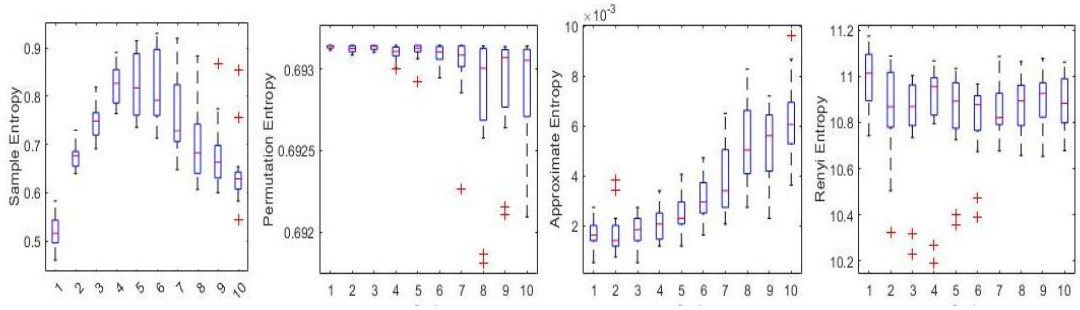
mance of ML classifiers, we used a ten-fold cross-validation technique and trained on Classification-Learner MATLAB tool. In this ten-fold cross-validation method, each event in our data is subjected to 10 iterations of testing. During training, each dataset is split into ten subsets, with one subset used for model validation and the remaining nine subsets forming a training subset for model assessment. To construct the model's performance measures, a portion of the dataset is used as the test set. The model's cross-validation performance is the arithmetic mean of the 10 performance estimates from the validation/test set.we used Baysen optimization as an optimizer, Expected improvement per second as acquisitionn function and iterate each classifier over 30 iterations.

This thesis studied the performance of a variety of machine learning classifiers through comprehensive experiments. As previously stated, this research employs fourteen machine learning classifiers (specified in section 3.6). Table 3.2 shows the parameter settings for various ML classifiers.

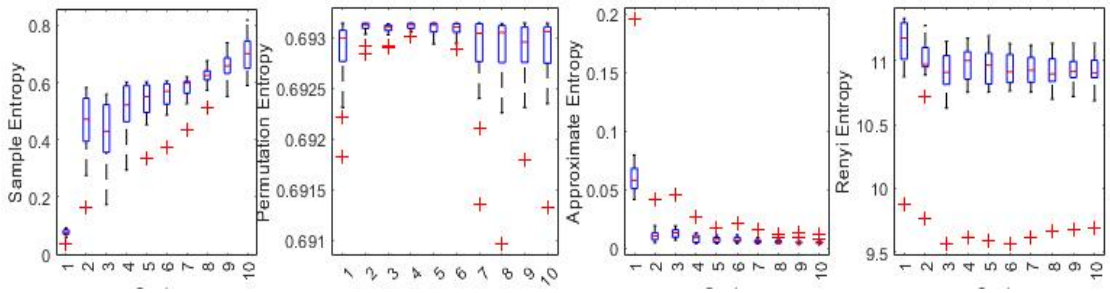
The proposed model generates four types of features/entropies using C2F and F2C MME calculation methods over 10 scales. The boxplots in Fig.4.2 exhibit the extracted sample entropy, permutation entropy, approximation entropy and renyi entropy feature sets for both HC and SZ classes. In this figure, the (a) and (c) represents the features set of HC class, while (b) and (d) represent the features



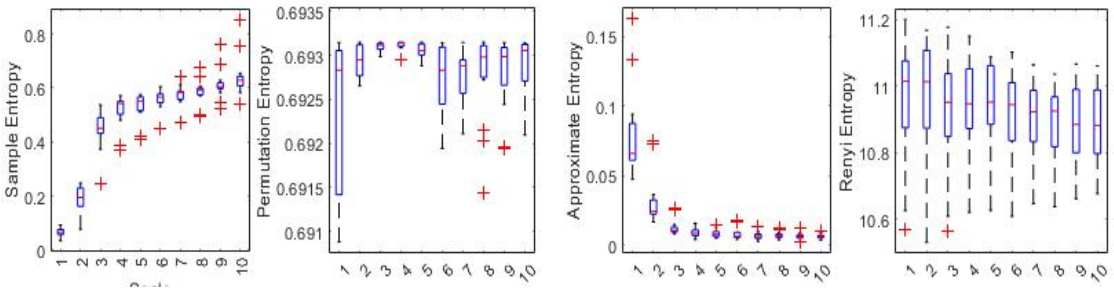
(a)



(b)



(c)



(d)

Figure 4.2: (a) Boxplots for different entropy features from one segment of HC class using the C2F MME method.(b) Boxplots for different entropy features from one segment of SZ class using the C2F MME method.(c) Boxplots for different entropy features from one segment of HC class using the F2C MME method.(d) Boxplots for different entropy features from one segment of SZ class using the F2C MME method.

set of SZ class and It is clear that the data distribution in both classes is identical. Fig.4.2 clearly shows the difference in values between the two groups, which supports effective classification.

4.2 Performance Assessment Parameters

This study computes the following assessment metrics, the ROC curve and the AUC, to analyse the performance of the proposed methodology:

$$\text{Acc (Accuracy)} = \frac{P^+ + P^-}{P^+ + P^- + p^+ + p^-} \times 100 \tag{4.1}$$

$$\text{Sen (Sensitivity)} = \frac{P^+}{P^+ + p^-} \times 100 \tag{4.2}$$

$$\text{Spe (Specificity)} = \frac{P^-}{P^- + p^+} \times 100 \tag{4.3}$$

$$\text{Pre(Precision)} = \frac{P^+}{P^+ + p^+} \times 100 \tag{4.4}$$

$$\text{F1 (F1-Score)} = \frac{2 * Pre * Sen}{Pre + Sen} \times 100 \tag{4.5}$$

$$\text{FPV (False Postive Rate)} = \frac{p^+}{P^- + p^+} \times 100 \tag{4.6}$$

$$\text{FNR (False Negative Rate)} = \frac{p^-}{P^+ + p^-} \times 100 \tag{4.7}$$

where

- (P^+): The patients with HC class and the proposed model also predict HC class.
- (P^-): The patients with SZ class and and the proposed model also predict SZ class.
- (p^+): The patients with SZ class, but the proposed model predicts HC class.
- (p^-): The patients with HC class, but the proposed model predicts SZ class.

4.3 Feature Selection Techniques

In this study, our proposed model generate 190 features for each 8 dataset. And we know that, every feature do not contribute equally in classification of data. For higher accuracy we are only take features with high contributing values. For feature selection we are using ANOVA and Kruskal-wallis techniques.

4.3.1 ANNOVA Test

Analysis of variance is a well-established mathematical process for comparing many different means. ANOVA also work as a tool for selecting features. [53]. There are several sample features for each of the n classes to be differentiated within the training set. The column of values at each frequency can be considered as an ANOVA-type problem with M observations on n classes in order to evaluate if the data provided at that frequency can be used to classify the sample as intended. The number generated by the equations is the "f-ratio" of the variance across classes to the variance within classes. The value of the f-ratio represents the degree of class difference. Data at retention periods that provide a f-ratio larger than a predetermined threshold are maintained as features, whilst all other data are removed. The variance across classes is calculated as follows:

$$\sigma_{er}^2 = \left(\sum\sum (x_{ij} - \bar{x_i})\right) - \left(\sum\sum (\bar{x_i} - \bar{x})\right)(M - n) \tag{4.8}$$

Where x_{ij} represents the i^{th} dimension of the j^{th} class. Then, a f-ratio is computed by dividing variances:

$$f_{ratio} = \sigma_{cl}^2 \sigma_{er}^2 \tag{4.9}$$

For each data point along with the feature set, or feature, an f-ratio is calculated. While the f-test ratio shows the degree of separation across classes, it does not tell the degree to which one class is isolated from the others. For example, if one class is significantly distanced from six others, the f-ratio may be greater than if all seven classes are only slightly apart. The selection of the f-ratio threshold value for feature selection was evaluated by calculating the sense of independence across classes on feature-selected data gathered using a range of threshold values. Once

an appropriate threshold value was calculated, the selected retention periods were separated from the remaining data for PCA-based pattern recognition analysis.

4.3.2 Kruskal-wallis Test

The Kruskal–Wallis statistic [54] is a non-parametric feature selection test.the Kruskal-Wallis method is used to identify relevant features, as it is comparatively less expensive and very easy to implement. The Kruskal-Wallis test examines whether or not two or more classes have the same median and returns the value P. Selecting attributes using discriminative information. If P is near to zero, the feature carries discriminatory information. otherwise, it will be removed. MME features are examined with the Kruskal-Wallis test. Features with test score less than threshold will be used for classification.

Chapter 5

RESULTS AND ANALYSIS

5.1 Results and Analysis

In this section, we show the test results for given method for multi-channel EEG signals from the database. The results of the study are based on bipolar classification. We used the HC and SC classes as binary classes. MVMD gives all the information in multi-channel EEG signals, which other decomposition methods can't control. MVMD take much computational time in comparison to other methods. The experiment is done on 8 datasets which include sample entropy, approximation entropy, renyi entropy and permutation entropy calculated by C2F and F2C MME method.

5.1.1 Results with F2C MME method

In this section, we are presenting the experiment result of the proposed methodology for fine-to-coarse based MME based feature extraction on different ML classifiers trained.table 5.1 represents the overall performance in form of F1-score (F1),sensitivity (Sen),precision (pre), accuracy (Acc), Roc-Auc (Auc), and specificity(Spc). In case of sample entropy, CSVM classifier achieved best/highest performance with 96.8% accuracy, 96.16% sensitivity, 98% Roc-Auc, 97.67% specificity, 96.9% f1-score and 98.04% precision. while ERT classifier achieved lowest performance with 87.9% accuracy, 87.06% sensitivity, 94% Roc-Auc, 88.95% specificity, 87.99% f1-score and 90.53% precision. In case of permutation entropy, EBAT classifier achieved best/highest performance with 77.8% accuracy, 83.96% sensitivity, 85% Roc-Auc, 78.48% specificity, 77.89% f1-score and 81.34% precision. While CKNN classifier

Table 5.1: Classification performance of F2C MME generated features with different ML classifiers

Entropies	**Sample Entropy**						**Permutation Entropy**						**Approximation Entropy.**						**Renyi Entropy**					
Classifiers	**Acc** (%)	**Sen** (%)	**Spe** (%)	**Pre** (%)	**F1** (%)	**Auc** (%)	**Acc** (%)	**Sen** (%)	**Spe** (%)	**Pre** (%)	**F1** (%)	**Auc** (%)	**Acc** (%)	**Sen** (%)	**Spe** (%)	**Pre** (%)	**F1** (%)	**Auc** (%)	**Acc** (%)	**Sen** (%)	**Spe** (%)	**Pre** (%)	**F1** (%)	**Auc** (%)
LR	89	92.08	90.89	92.08	91.48	92	73.3	75.23	70.93	75.84	73.01	82	72.7	75.37	69.57	74.38	72.35	77	71.4	73.71	68.6	73.35	71.06	79
LDA	90.7	87.69	94.37	94.98	90.9	96	69.1	71.08	66.66	72.12	68.79	76	69	74.71	62.4	69.96	68	73	69.9	72.72	66.66	71.89	69.55	78
QDA	89.3	85.62	93.79	94.36	89.51	94	70.1	63.09	78.68	78.12	70.02	77	71.2	61.98	81.97	80.12	70.58	80	72.8	77.19	67.63	73.65	72.09	81
LSVM	89.1	87.53	90.89	92.1	89.17	95	72.2	70.12	74.61	77.01	72.29	80	76.2	88.42	61.82	73.08	72.76	82	71.9	71.23	72.67	75.34	71.94	80
QSVM	95.6	94.88	96.51	97.05	95.68	99	75.9	73.32	79.06	80.95	76.08	84	81.9	84.13	79.26	82.62	81.62	87	75.9	78.84	72.48	77.05	75.52	83
CSVM	96.8	96.16	97.67	98.04	96.9	98	75	75.39	74.61	78.27	74.99	82	80	87.1	71.7	78.3	78.65	89	77.5	79.33	75.38	79.07	77.3	84
FKNN	94.3	94.88	93.6	94.73	94.23	94	67.4	68.84	65.69	70.88	67.22	67	82.7	82.31	83.13	85.12	82.71	83	74.4	79.33	68.6	74.76	73.57	74
CKNN	89.3	89.13	89.53	91.17	89.32	96	62.9	65.49	59.68	66.34	62.45	70	76.4	75.37	77.51	79.72	76.42	85	71.8	74.21	68.99	73.72	71.5	80
WKNN	91.9	91.69	92.05	93.33	91.86	97	69.4	65.01	74.8	75.79	69.56	77	82.6	84.13	80.81	83.71	82.43	91	75.9	84.79	65.5	74.24	73.9	82
EBT	92.2	92.01	92.44	93.65	92.22	98	75.7	72.84	79.26	80.99	75.91	86	80.4	77.35	83.91	84.93	80.49	89	70.6	71.23	69.76	73.42	70.48	79
EBAT	91.9	91.37	92.63	93.77	91.99	97	77.8	77.31	78.48	81.34	77.89	85	82.1	80.99	83.33	85.06	82.14	90	74.9	72.56	77.41	79.24	75.04	81
ESD	91.2	89.29	93.41	94.26	91.3	97	72.9	75.87	69.37	75.03	72.47	80	72.3	81.15	61.82	71.36	70.17	80	73.2	75.53	70.54	75.04	72.94	82
ESK	95.1	95.52	94.57	95.52	95.04	98	75.6	71.88	80.03	81.37	75.73	85	83.8	83.96	83.52	85.66	83.73	92	74.6	80.49	67.63	74.46	73.5	82
ERT	87.9	87.06	88.95	90.53	87.99	94	70.4	61.98	80.62	79.5	70.08	80	74.9	69.91	80.81	81.03	74.96	86	68	62.64	74.22	74.02	67.94	76

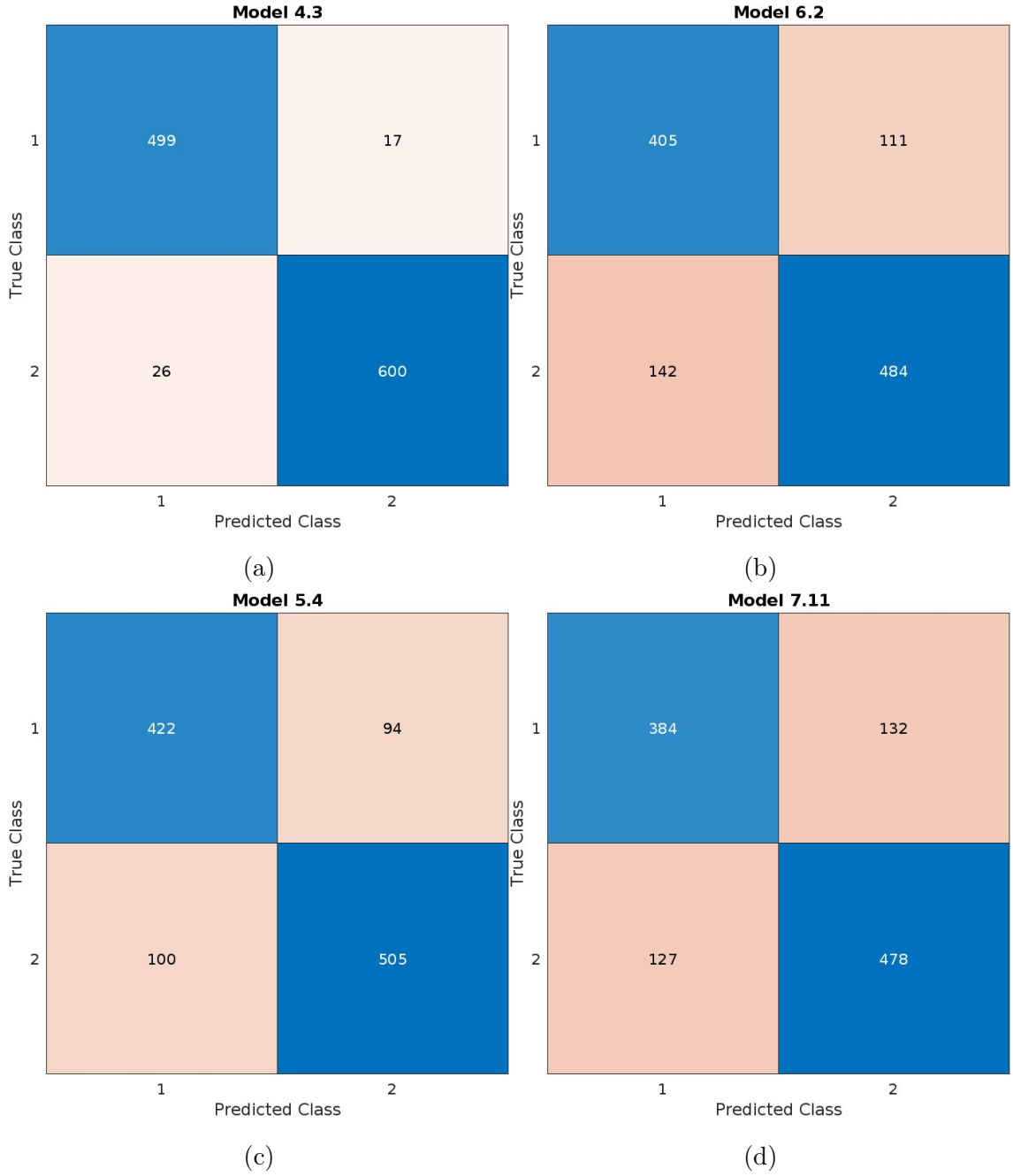


Figure 5.1: Confusion matrix of (a) CSVM classifier on F2C MME based sample entropy data. (b) EBAT classifier on F2C MME based permutation entropy data. (c)ESK classifier on F2C MME based approximation entropy data. (d) CSVM classifier on F2C MME based renyi entropy data.

achieved lowest performance with 62.9% accuracy, 65.49% sensitivity, 70% Roc-Auc, 59.68% specificity, 62.45% f1-score and 62.45% precision. In case of approx entropy, ESK classifier achieved best/highest performance with 83.8% accuracy, 83.96% sensitivity, 92% Roc-Auc, 83.52% specificity, 83.73% f1-score and 85.66% precision.

While LDA classifier achieved lowest performance with 69% accuracy, 74.41% sensitivity, 73% Roc-Auc, 62.4% specificity, 68% f1-score and 69.96% precision. In case of renyi entropy, CSVM classifier achieved best/highest performance with 77.5% accuracy, 79.33% sensitivity, 84% Roc-Auc, 75.38% specificity, 77.3% f1-score and 79.07% precision. While ERT classifier achieved lowest performance with 68% accuracy, 62.64% sensitivity, 76% Roc-Auc, 74.22% specificity, 67.94% f1-score and 74.02% precision. From the results, we can say that, sample entropy provides best results in comparison to other entropies. And their performance order is *Sample entropy* > *Approx entropy* > *Permutation entropy* > *Renyi Entropy*. But in case of worst performance this order will become *Sample entropy* > *Approx entropy* > *Renyi Entropy* > *Permutation entropy*.

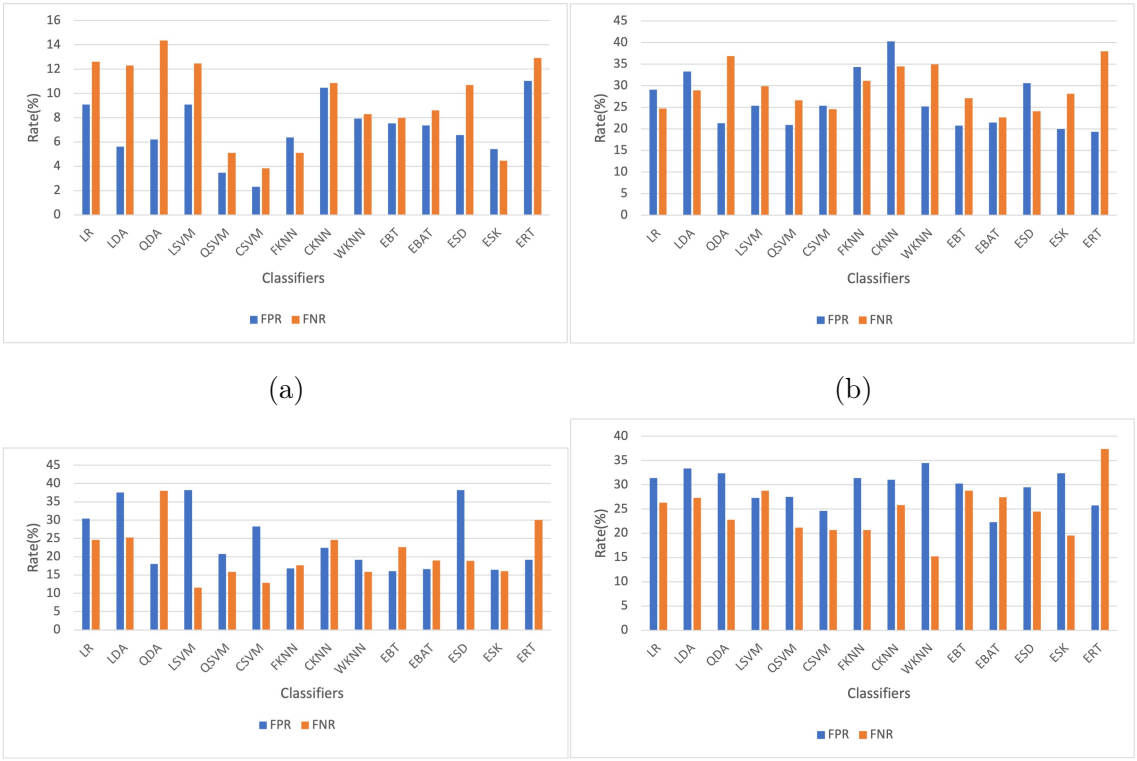


Figure 5.2: (a) Comparison of FPR and FNR on sample entropy data. (b) Comparison of FPR and FNR on permutation entropy data. (c) Comparison of FPR and FNR on approximation entropy data. (d) Comparison of FPR and FNR on F2C MME based renyi entropy data.

Fig.5.2 shows a confusion matrix for each of the top performing classifiers utilising a ten - fold cross-validation approach on their F2C MME-based respected entropies

dataset for more thorough performance verification, such as class-specific performance. these confusion matrix obtained on 20% testdata (subset of dataset). in figure 5.1 rows represent predicted class and column represent true class.Figure 5.2 represents a comparison between FPR and FNR for each F2C MME-based entropy in order to provide information on classification efficiency. if FPR and FNR are low, model will be more efficient. In case of sample entropy low FPR(2.32) and FNR(3.83) are achieved by CSVM classifier. while ERT classifier achieve worst FPR(11.04) and FNR(12.93). In case of permutation entropy low FPR(21.51) and FNR(22.68) are achieved by EBAT classifier. while CKNN classifier achieve worst FPR(40.31) and FNR(25.28). In case of approximation entropy low FPR(16.47) and FNR(16.03) are achieved by ESK classifier. while LDA classifier achieve worst FPR(37.59) and FNR(25.28). In case of renyi entropy low FPR(24.61) and FNR(20.66) are achieved by CSVM classifier. while ERT classifier achieve worst FPR(25.77) and FNR(37.35).

5.1.2 Results with C2F MME method

In this section, we are presenting the experiment result of the proposed methodology for coarse-to-fine based MME based feature extraction on different ML classifiers. Table 5.2 represents the overall performance in form of F1-score (F1),sensitivity (Sen), precision (pre),accuracy (Acc), Roc-Auc (Auc), and specificity(Spc).

In case of sample entropy, ESK and FKNN classifiers achieved best/highest performance with 97.1% accuracy, 97.12% sensitivity, 97.09% Roc-Auc, 97.09% specificity, 97.1% f1-score and 97.59% precision. While ERT classifier achieved lowest performance with 89.7% accuracy, 86.42% sensitivity, 96% Roc-Auc, 93.6% specificity, 89.86% f1-score and 94.35% precision. In case of permutation entropy, QSVM classifier achieved best/highest performance with 82.2% accuracy, 76.51% sensitivity, 91% Roc-Auc, 89.14% specificity, 82.34% f1-score and 89.53% precision. While LDA classifier achieved lowest performance with 65.1% accuracy, 62.61% sensitivity, 71% Roc-Auc, 68.21% specificity, 68% f1-score and 70.5% precision.In case of approx entropy, ESK classifier achieved best/highest performance with 86.9% accuracy, 83.14% sensitivity, 94% Roc-Auc, 91.27% specificity, 87.01% f1-score and 91.78% precision. While LDA classifier achieved lowest performance with 69% accuracy, 74.41% sensitivity, 73% Roc-Auc, 62.4% specificity, 68% f1-score and 69.96%

Table 5.2: Classification performance of C2F MME generated features with different ML classifiers

Entropies	**Sample Entropy**						**Permutation Entropy**						**Approximation Entropy.**						**Renyi Entropy**					
Classifiers	**Acc** (%)	**Sen** (%)	**Spe** (%)	**Pre** (%)	**F1** (%)	**Auc** (%)	**Acc** (%)	**Sen** (%)	**Spe** (%)	**Pre** (%)	**F1** (%)	**Auc** (%)	**Acc** (%)	**Sen** (%)	**Spe** (%)	**Pre** (%)	**F1** (%)	**Auc** (%)	**Acc** (%)	**Sen** (%)	**Spe** (%)	**Pre** (%)	**F1** (%)	**Auc** (%)
LR	92.3	91.85	92.82	93.95	92.33	77	76.6	78.27	74.61	78.9	76.39	80	72.7	75.37	69.57	74.38	72.35	77	70.4	72.39	68.02	72.63	70.13	77
LDA	93	92.65	93.41	94.46	93.02	73	65.1	62.61	68.21	70.5	65.29	71	69	74.71	62.4	69.96	68	73	69.3	74.04	63.75	70.55	68.51	75
QDA	95.3	96.8	93.41	94.68	95.07	80	76.9	62.46	94.37	93.09	75.16	79	71.2	61.98	81.97	80.12	70.58	80	77.7	71.57	84.88	84.73	77.65	85
LSVM	87.3	86.1	88.75	90.28	87.4	82	79.5	67.57	93.99	93.17	78.61	89	76.2	88.42	61.82	73.08	72.76	82	73.6	79.66	66.47	73.58	72.46	80
QSVM	95.6	94.88	96.51	97.05	95.68	87	82.2	76.51	89.14	89.53	82.34	91	81.9	84.13	79.26	82.62	81.62	87	79.9	79	81	82.98	79.98	86
CSVM	96.3	95.68	97.09	97.55	96.37	89	79.7	66.45	95.73	94.97	78.44	89	80	87.1	71.7	78.3	78.65	89	82.9	81.65	84.3	85.91	82.95	88
FKNN	97.1	97.6	96.51	94.13	97.05	83	78.4	80.19	76.16	80.32	78.12	78	82.7	82.31	83.13	85.12	82.71	83	81.6	83.3	79.65	82.75	81.43	81
CKNN	93.3	93.61	93.02	94.21	93.31	85	79.2	76.67	82.36	84.06	79.41	88	76.4	75.37	77.51	79.72	76.42	85	79.3	74.38	85.85	86.04	79.7	88
WKNN	96	95.68	96.31	96.92	95.99	91	80.4	75.55	86.24	86.94	80.54	88	82.6	84.13	80.81	83.71	82.43	91	83.4	83.22	83.52	85.49	83.36	92
EBT	94.6	93.76	95.54	96.22	94.64	89	80	73.48	87.98	88.12	80.07	88	80.4	77.35	83.91	84.93	80.49	89	79.3	73.38	86.24	86.21	79.29	87
EBAT	94.1	93.61	94.76	95.59	94.18	90	82	74.6	90.89	90.85	81.94	89	82.1	80.99	83.33	85.06	82.14	90	80.1	74.87	86.24	86.45	80.15	88
ESD	92.8	91.69	94.18	95.03	92.91	80	68.9	69.16	68.6	72.77	68.87	76	72.3	81.15	61.82	71.36	70.17	80	73.4	78.34	67.63	73.94	72.59	80
ESK	97.1	97.12	97.09	97.59	97.1	92	81.9	79.07	85.27	86.69	82.05	91	86.9	83.14	91.27	91.78	87.01	94	84.8	84.79	84.88	86.8	84.83	92
ERT	89.7	86.42	93.6	94.25	89.86	86	78.9	66.77	93.6	92.68	77.94	87	74.9	69.91	80.81	81.03	74.96	86	77.2	69.42	86.24	85.53	76.92	86

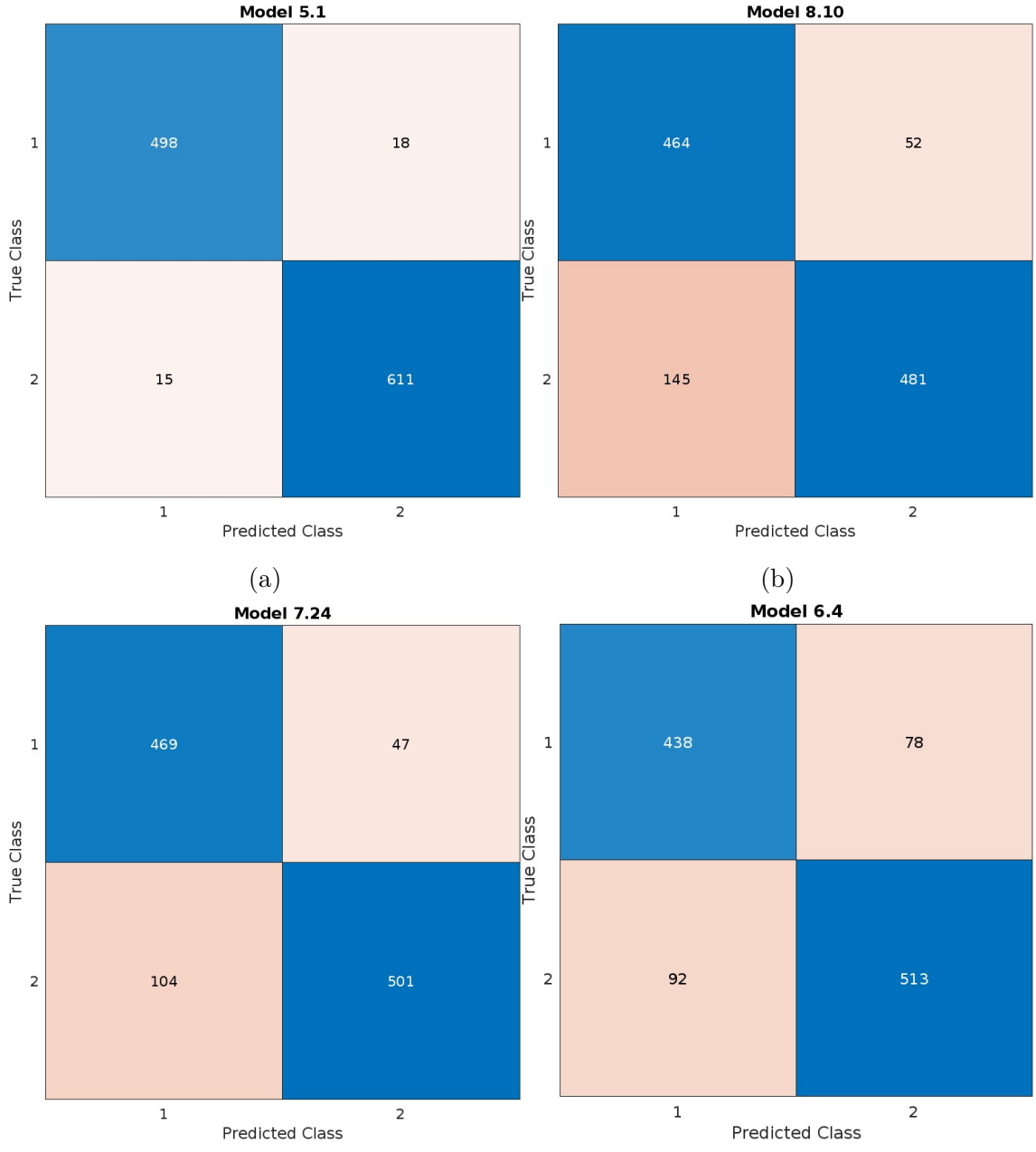


Figure 5.3: Confusion matrix of (a)ESK classifier on sample entropy data. (b)QSVM classifier on permutation entropy data. (c) ESK classifier on approximation entropy data. (d)ESK classifier on C2F MME based renyi entropy data.

precision. In case of renyi entropy, ESK classifier achieved best/highest performance with 84.8% accuracy, 84.79% sensitivity, 92% Roc-Auc, 84.88% specificity, 84.33% f1-score and 86.8% precision. While LDA classifier achieved lowest performance with 69.3% accuracy, 74.04% sensitivity, 75% Roc-Auc, 63.75% specificity, 68.51%

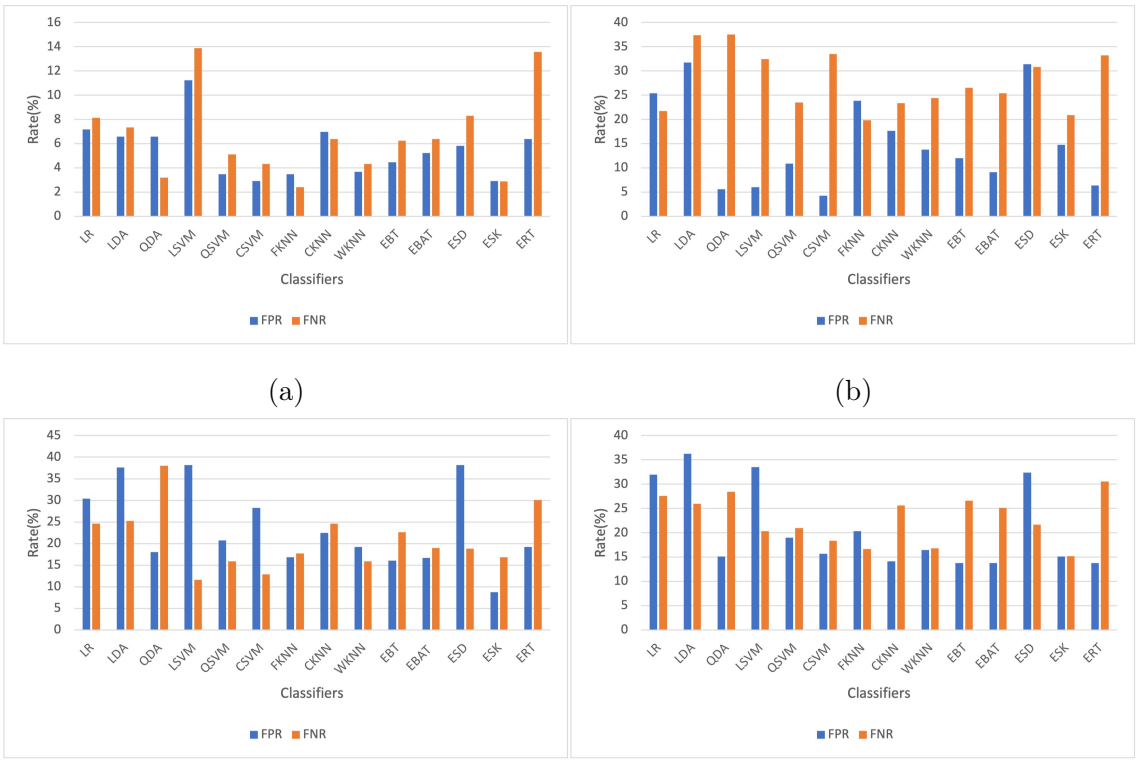


Figure 5.4: (a) Comparison between FPR nad FNR on sample entropy data. (b) Comparison between FPR nad FNR on permutation entropy data. (c) Comparison between FPR nad FNR on approximation entropy data. (d) Comparison between FPR nad FNR on C2F MME based renyi entropy data.

f1-score and 70.55% precision. From the results, we can say that, sample entropy provides best results in comparison to other entropies. and their performance order is *Sample entropy* > *Approx entropy* > *Renyi Entropy* > *Permutation entropy* for overall best results. But in case of worst performance this order will become *Sample entropy* > *Renyi Entropy* > *Approx entropy* > *Permutation entropy*.

Fig.5.4 shows a confusion matrix for each of the top performing classifiers utilising a ten-fold cross-validation approach on their F2C MME-based respected entropies dataset for more thorough performance verification, such as class-specific performance. Figure 5.4 represents a comparison of FPR and FNR for each F2C MME-based entropy in order to provide information on classification efficiency. In case of sample entropy low FPR(2.9) and FNR(2.87) are achieved by ESK classifier. While ERT classifier achieve worst FPR(6.36) and FNR(13.57). In case of permutation entropy low FPR(10.85) and FNR(23.48) are achieved by QSVM classifier.

While LDA classifier achieve worst FPR(31.78) and FNR(37,38). In case of approximation entropy low FPR(8.72) and FNR(16.85) are achieved by ESK classifier. While LDA classifier achieve worst FPR(37.59) and FNR(25.28). In case of renyi entropy low FPR(15.11) and FNR(15.2) are achieved by ESK classifier. While ERT classifier achieve worst FPR(36.24) and FNR(25.95).

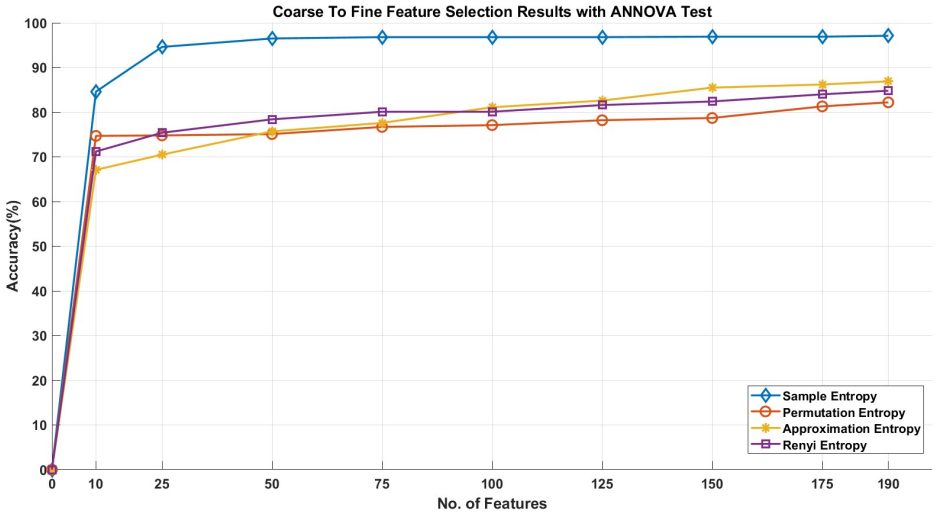
5.2 Result Comparison

we are using ANNNOVA test and Kruskal-wallis test to show, how accuracy varies with number of features for different entropies. In case of C2F MME entropies Figure 5.5 shows, for sample entropy Kruskal-wallis test provide better results in comparison to ANNOVA test. Where accuracy is enhanced from 97.1% to 97.2%. In case of permutation entropy,approx entropy and renyi entropy ANNOVA test shows better results. But it do not increase accuracy.

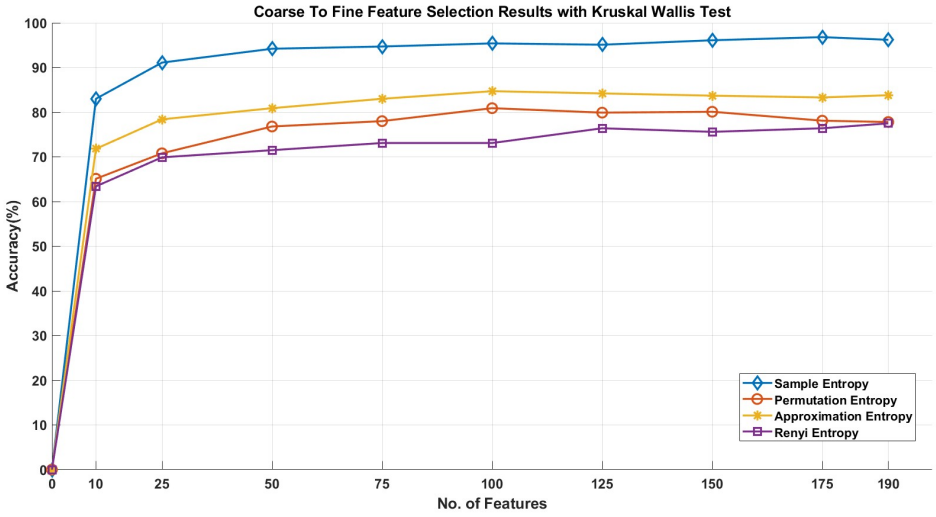
In case of F2C MME entropies Figure 5.6 shows, for sample entropy ANNOVA test provide better results in comparion to kruskal-wallis test. Where accuracy is enhanced from 97.2% to 96.8%. In case of permutation entropy, both test enhance accuracy from 77.8% to 80.1%. In case of renyi entropy ANNOVA test increase accuracy from 77.5% to 76.4% and Kruskal-wallis test increase accuracy from 77.5% to 77.3%. In case of approximation entropy ANNOVA test increase accuracy from 83.8% to 84.7% and Kruskal-wallis test increase accuracy from 83.8% to 86.9%. According to Figure 5.6 as number of feature increase, accuracy also increase. But it became constant after a certain number of features.

Table 5.3: Comparing proposed approach results with existing Schizophrenia detection models.

Authors	**Techniques**	**Classifiers**	**Performance**
Bilas et al. [19] (2021)	Multivariate iterative filtering	SVM(cubic)	Acc = 98.9%,Sen = 99.0%, Spe = 98.8%,AUC = 99.9%
Buettner et al. [27] (2020)	FFT	RF	Acc = 96.77%,Sen = 96.77%,Spe = 96.77%
Khare et al. [21](2021)	SPWVD TF plots	CNN	Acc = 93.36%,Sen = 94.25%,Spe = 92.05%
Siuly et al. [23]	EWT	EBAT	Acc = 93.21%,Sen = 93.20%,Spe = 84.33%
Shalbaf et al. [26](2020)	CWT	ResNet18-SVM	Acc = 98.60%,Sen = 99.65%,Spe = 96.92%
Krishananet al. [29]	MEMD based entropies	SVM(RBF)	Acc = 96.3%,Sen = 94%,Spe = 92%
This work	MME based entropies (sample entropy)	ESK	Acc = 97.1%,Sen = 97.12%,Spe = 97.09%,AUC = 98%

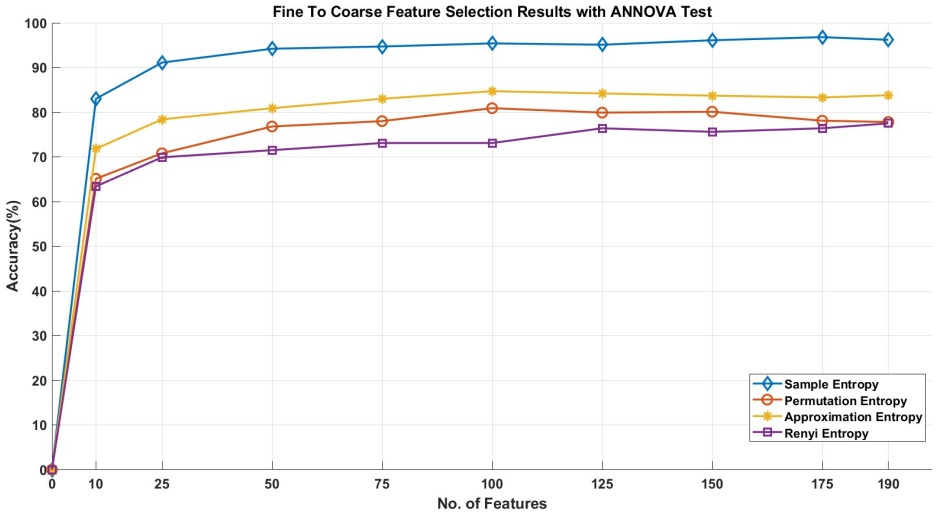


(a)

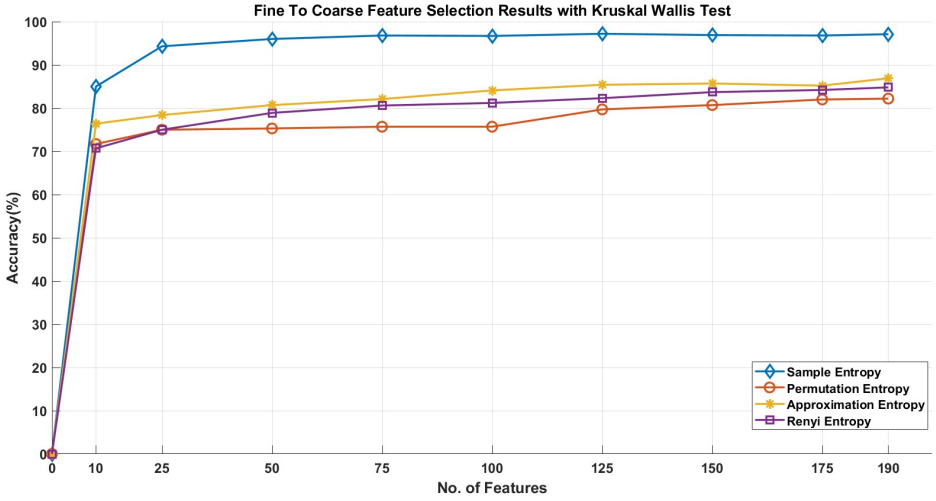


(b)

Figure 5.5: Accuracy comparison with the different number of features on (a)C2F MME based entropies with ANOVA test. (b) C2F MME based entropies with Kruskal-Wallis test.



(a)



(b)

Figure 5.6: Accuracy comparison with the different number of features on (a) F2C MME based entropies with ANOVA test. (b) F2C MME based entropies with Kruskal-Wallis test.

The performance of the suggested methodology was compared with existing Schizophrenia detection approach, and the results are displayed in table 5.3. The table shows that proposed approach perform well in comparison to with recently suggested Schizophrenia detection methods.

5.3 Overview

This study found that reciprocal information between multichannel EEG signals is also necessary for the identification of schizophrenia. This study shows that high frequency plays a crucial role in the diagnosis of Schizophrenia. In the case of coarse-to-fine sample entropy, we get the highest classification performance. The proposed method with sample entropy is a better feature extraction method in comparison to other entropies, as it achieves maximum classification throughout the study. And for classifiers, ensemble classifiers get good performance in comparison to others. We can conclude that the multi-variate multiscale sample entropy of MMOs produced by the MVMD approach is a perfect representation of multi-channel EEG signals for the diagnosis of Schizophrenia. And in general ESK is best choice for Schizophrenia detection.

Chapter 6

CONCLUSION

6.1 Conclusion

In this study, a multivariate multiscale entropy-based feature calculation method is proposed for automatically recognising Schizophrenia from multichannel EEG signals. We used MVMD decomposition to extract MMOs from an EEG signal in proposed methodology. Further we used coarse-to-fine based multivariate multiscale entropy calculation method to generate features from these MMOs. For efficient results, different type of entropies: sample entropy, approximation entropy, permutation entropy and renyi entropy are calculated by MMSE, which generating different type of features. These generated features is feed to fourteen ML classifiers. For detailed classification performance, we calculated Acc, Spe, Sen, F1-score, pre and ROC-AUC of each classifiers. The proposed method is studied on publicly available dataset of Schizophrenia. In comparison to recent developed Schizophrenia recognition methods our proposed method outperform many techniques. We can say that our proposed approach is good enough to recognise Schizophrenia and Healthy Control from multichannel EEG signals efficiently.

6.2 Future Work

In the future, we try to adopt a low computation cost, real-time and more efficient method for Schizophrenia diagnosis. we will try to make a portable system using this approach to diagnosis Schizophrenia in real time.

REFERENCES

[1] *Schizophrenia.*
https://www.who.int/news-room/fact-sheets/detail/schizophrenia.

[2] *What is Schizophrenia & its symptoms?*
https://psychiatry.org/patients-families/schizophrenia/what-is-schizophrenia.

[3] *Treatment for Schizophrenia.*
https://www.webmd.com/schizophrenia/mental-health-schizophrenia.

[4] T. H. McGlashan, "Early detection and intervention of schizophrenia: rationale and research," *The British Journal of Psychiatry*, vol. 172, no. S33, pp. 3–6, 1998.

[5] T. M. Laursen, M. Nordentoft, and P. B. Mortensen, "Excess early mortality in schizophrenia," *Annual review of clinical psychology*, vol. 10, pp. 425–448, 2014.

[6] D. Sulzer, C. Cassidy, G. Horga, U. J. Kang, S. Fahn, L. Casella, G. Pezzoli, J. Langley, X. P. Hu, F. A. Zucca *et al.*, "Neuromelanin detection by magnetic resonance imaging (mri) and its promise as a biomarker for parkinson's disease," *NPJ Parkinson's disease*, vol. 4, no. 1, pp. 1–13, 2018.

[7] U. R. Acharya, S. L. Fernandes, J. E. WeiKoh, E. J. Ciaccio, M. K. M. Fabell, U. J. Tanik, V. Rajinikanth, and C. H. Yeong, "Automated detection of alzheimer's disease using brain mri images–a study with various feature extraction techniques," *Journal of Medical Systems*, vol. 43, no. 9, pp. 1–14, 2019.

[8] N. Salamon, J. Kung, S. Shaw, J. Koo, S. Koh, J. Wu, J. Lerner, R. Sankar, W. Shields, J. Engel *et al.*, "Fdg-pet/mri coregistration improves detection of cortical dysplasia in patients with epilepsy," *Neurology*, vol. 71, no. 20, pp. 1594–1601, 2008.

[9] V. Gupta and R. B. Pachori, "Epileptic seizure identification using entropy of fbse based eeg rhythms," *Biomedical Signal Processing and Control*, vol. 53, p. 101569, 2019.

[10] V. Joshi, R. B. Pachori, and A. Vijesh, "Classification of ictal and seizure-free eeg signals using fractional linear prediction," *Biomedical Signal Processing and Control*, vol. 9, pp. 1–5, 2014.

[11] A. Bhattacharyya and R. B. Pachori, "A multivariate approach for patient-specific eeg seizure detection using empirical wavelet transform," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2003–2015, 2017.

[12] R. B. Pachori and V. Gupta, "Biomedical engineering fundamentals," in *Intelligent Internet of Things*. Springer, 2020, pp. 547–605.

[13] V. Jahmunah, S. L. Oh, V. Rajinikanth, E. J. Ciaccio, K. H. Cheong, N. Arunkumar, and U. R. Acharya, "Automated detection of schizophrenia using nonlinear signal processing methods," *Artificial intelligence in medicine*, vol. 100, p. 101698, 2019.

[14] A. Khosla, P. Khandnor, and T. Chand, "A comparative analysis of signal processing and classification methods for different applications based on eeg signals," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 2, pp. 649–690, 2020.

[15] A. Goshvarpour and A. Goshvarpour, "Schizophrenia diagnosis using innovative eeg feature-level fusion schemes," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 1, pp. 227–238, 2020.

[16] J. M. Ford, V. A. Palzes, B. J. Roach, and D. H. Mathalon, "Did i do that? abnormal predictive processes in schizophrenia when button pressing to deliver a tone," *Schizophrenia bulletin*, vol. 40, no. 4, pp. 804–812, 2014.

[17] N. N. Boutros, C. Arfken, S. Galderisi, J. Warrick, G. Pratt, and W. Iacono, "The status of spectral eeg abnormality as a diagnostic test for schizophrenia," *Schizophrenia research*, vol. 99, no. 1-3, pp. 225–237, 2008.

[18] B. J. Roach and D. H. Mathalon, "Event-related eeg time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia," *Schizophrenia bulletin*, vol. 34, no. 5, pp. 907–926, 2008.

[19] K. Das and R. B. Pachori, "Schizophrenia detection technique using multivariate iterative filtering and multichannel eeg signals," *Biomedical Signal Processing and Control*, vol. 67, p. 102525, 2021.

[20] R. Buettner, M. Hirschmiller, K. Schlosser, M. Rössle, M. Fernandes, and I. J. Timm, "High-performance exclusion of schizophrenia using a novel machine learning method on eeg data," in *2019 IEEE International Conference on E-Health Networking, Application & Services (HealthCom)*. IEEE, 2019, pp. 1–6.

[21] S. K. Khare, V. Bajaj, and U. R. Acharya, "Spwvd-cnn for automated detection of schizophrenia patients using eeg signals," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–9, 2021.

[22] A. Shoeibi, D. Sadeghi, P. Moridian, N. Ghassemi, J. Heras, R. Alizadehsani, A. Khadem, Y. Kong, S. Nahavandi, Y.-D. Zhang *et al.*, "Automatic diagnosis of schizophrenia in eeg signals using cnn-lstm models," *Frontiers in Neuroinformatics*, vol. 15, 2021.

[23] S. Siuly, S. K. Khare, V. Bajaj, H. Wang, and Y. Zhang, "A computerized method for automatic detection of schizophrenia using eeg signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 11, pp. 2390–2400, 2020.

[24] Z. Aslan and M. Akın, "Automatic detection of schizophrenia by applying deep learning over spectrogram images of eeg signals," 2020.

[25] S. L. Oh, J. Vicnesh, E. J. Ciaccio, R. Yuvaraj, and U. R. Acharya, "Deep convolutional neural network model for automated diagnosis of schizophrenia using eeg signals," *Applied Sciences*, vol. 9, no. 14, p. 2870, 2019.

[26] A. Shalbaf, S. Bagherzadeh, and A. Maghsoudi, "Transfer learning with deep convolutional neural network for automated detection of schizophrenia from eeg signals," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 4, pp. 1229–1239, 2020.

[27] R. Buettner, D. Beil, S. Scholtz, and A. Djemai, "Development of a machine learning based algorithm to accurately detect schizophrenia based on one-minute eeg recordings," in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020.

[28] M. Sabeti, S. Katebi, and R. Boostani, "Entropy and complexity measures for eeg signal classification of schizophrenic and control participants," *Artificial intelligence in medicine*, vol. 47, no. 3, pp. 263–274, 2009.

[29] P. T. Krishnan, A. N. J. Raj, P. Balasubramanian, and Y. Chen, "Schizophrenia detection using multivariateempirical mode decomposition and entropy measures from multichannel eeg signal," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 3, pp. 1124–1139, 2020.

[30] N. ur Rehman and H. Aftab, "Multivariate variational mode decomposition," *IEEE Transactions on signal processing*, vol. 67, no. 23, pp. 6039–6052, 2019.

[31] E. Olejarczyk and W. Jernajczyk, "Eeg in schizophrenia," 2017. [Online]. Available: https://doi.org/10.18150/repod.0107441

[32] ——, "Graph-based analysis of brain connectivity in schizophrenia," *PloS one*, vol. 12, no. 11, p. e0188629, 2017.

[33] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE transactions on signal processing*, vol. 62, no. 3, pp. 531–544, 2013.

[34] G. Baldini, "On the application of entropy measures with sliding window for intrusion detection in automotive in-vehicle networks," *Entropy*, vol. 22, no. 9, p. 1044, 2020.

[35] P. Jizba and T. Arimitsu, "Observability of rényi's entropy," *Phys. Rev. E*, vol. 69, p. 026128, Feb 2004. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.69.026128

[36] G. Wenye, *Shannon and Non-Extensive Entropy, MATLAB Central File Exchange.*, 2020. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/18133-shannon-and-nonextensive-entropy

[37] S. M. Pincus, "Approximate entropy as a measure of system complexity." *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.

[38] B. D. Fulcher and N. S. Jones, "hctsa: A computational framework for automated time-series phenotyping using massive feature extraction," *Cell systems*, vol. 5, no. 5, pp. 527–531, 2017.

[39] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical review letters*, vol. 88, no. 17, p. 174102, 2002.

[40] H. Azami, M. Rostaghi, D. Abásolo, and J. Escudero, "Refined composite multiscale dispersion entropy and its application to biomedical signals," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 12, pp. 2872–2879, 2017.

[41] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, 2000.

[42] H. Azami, A. Fernández, and J. Escudero, "Refined multiscale fuzzy entropy based on standard deviation for biomedical signal analysis," *Medical & biological engineering & computing*, vol. 55, no. 11, pp. 2037–2052, 2017.

[43] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of biological signals," *Physical review E*, vol. 71, no. 2, p. 021906, 2005.

[44] M. Hu and H. Liang, "Adaptive multiscale entropy analysis of multivariate neural data," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 1, pp. 12–15, 2011.

[45] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

[46] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.

[47] H. Drucker and C. Cortes, "Boosting decision trees," *Advances in neural information processing systems*, vol. 8, 1995.

[48] T. G. Dietterich, "An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization," *Machine learning*, vol. 40, no. 2, pp. 139–157, 2000.

[49] J. Hamm and D. D. Lee, "Grassmann discriminant analysis: a unifying view on subspace-based learning," in *Proceedings of the 25th international conference on Machine learning*, 2008, pp. 376–383.

[50] T. K. Ho, "Nearest neighbors in random subspaces," in *Joint IAPR international workshops on statistical techniques in pattern recognition (SPR) and structural and syntactic pattern recognition (SSPR)*. Springer, 1998, pp. 640–648.

[51] S. Liu, S. You, Z. Lin, C. Zeng, H. Li, W. Wang, X. Hu, and Y. Liu, "Data-driven event identification in the us power systems based on 2d-olpp and rusboosted trees," *IEEE Transactions on Power Systems*, vol. 37, no. 1, pp. 94–105, 2021.

[52] V. Padhmashree and A. Bhattacharyya, "Human emotion recognition based on time–frequency analysis of multivariate eeg signal," *Knowledge-Based Systems*, vol. 238, p. 107867, 2022.

[53] K. J. Johnson and R. E. Synovec, "Pattern recognition of jet fuels: comprehensive gc× gc with anova-based feature selection and principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 60, no. 1-2, pp. 225–237, 2002.

[54] S. Vora and H. Yang, "A comprehensive study of eleven feature selection algorithms and their impact on text classification," in *2017 Computing Conference*. IEEE, 2017, pp. 440–449.