Machine Learning-Based Approach for Identification of Seizure Onset Zone in Patients with Focal Epilepsy from Interictal Intracranial EEG

PH.D. THESIS

MOST. SHEULI AKTER

A thesis submitted in fulfillment of the requirements

for the degree of Doctor of Philosophy

Department of Electronic and Information Engineering

Tokyo University of Agriculture and Technology

February 14, 2021

Abstract

Epilepsy is a chronic disease of the central nervous system characterized by repeated and unpredictable seizure. Approximately, 1 million people in Japan (Japan epilepsy society) and 50 million people all over the world (WHO) are diagnosed as epilepsy.

A seizure is a sudden disturbance in the brain's neural activity that produces disruptive physical prefix such as a lapse in attention and memory, a sensory hallucination, or an entire-body convulsion. Roughly, 1 out of every 3 individuals with epilepsy continues to experience with intractable seizures which is also known as drug registant epilepsy. To control seizures with intractable epilepsy, surgical treatment may, the best possible solution for seizure freedom. Through surgical treatment, the epileptologist remove that area of cortex from where the seizures are initiated, which known as the seizure onset zone (SOZ). To localize the SOZ electrodes from a part of the irritable and symptomatic zone, it is necessity for epileptologist's to observe the long term multichannel iEEG data by visual inspection, which is very time consuming and laborious process. Moreover, there is a shortage of clinical experts for such diagnosis.

Therefore designing a machine learning based computer aided solution for localizing the seizure onset zone (SOZ) in patients with focal epilepsy from interictal and ictal iEEG has been expected to epileptologists. The aim of this thesis is to detect the SOZ electrodes using high frequency components (HFC) with machine learning approach. It is known that, the high frequency components (>80 Hz) including ripple and fast ripple bands, of interictal iEEG signal, are associated to the epileptic seizure. In this thesis,

our proposed computer aided solution provides an intuition for epileptologist which assist them in two ways: (1) to observe the localization of SOZ and non-SOZ segments over duration of the iEEG data, and (2) gives feasible information about the active electrodes located close to the SOZ electrodes.

For localization of SOZ electrodes, in our first proposed design, we have used entropy based feature extraction methods and radial basis function kernel-based SVM classifier with 10-fold cross validation. In this solution, to select the prominent entropy features, we have used sparse linear discriminant analysis (sLDA). Due to the imbalance of SOZ and non-SOZ electrodes in iEEG data, the usage of machine learning techniques is always tricky. To handle this problem, we have used an adaptive synthetic oversampling approach (ADASYN) in the training stage of SVM classifier. Eight patients were examined to observe the efficiency of the proposed design.

However, in the above methodology, firstly, the performance of entropy estimation strongly depends on the appropriate parameter selection. Besides, some entropy features have higher computational cost. Secondly, the detection of SOZ channels were determined based on the number of detected segments on SOZ and non-SOZ channels and the detection decision was made by SVM classifier based hard thresholding. Thirdly, epileptic activities related appropriate bands selection not reported in the previous studies that may improve the performance of computer-aided solution. To address these problems, in our second proposed framework, we have used statistical features with information theoretic entropy features for identifying SOZ electrodes. We hypothesis that, these statistical features are effective to characterize the epileptic activities in iEEG signal, since these features have already been used in other context of epilepsy studies. Mutual information (MI) scores based data-driven grid-search method was developed to jointly optimize the bands and features. A LightGBM classifier was used to score each segment of a channels and final score of a channel was obtained by

averaging the scores of all segments of that channel. The probable SOZ channels were localized based on the higher scores of the channels. To observe the efficiency of method, we used eleven patients with medically intractable epilepsy caused by focal cortical dysplasia (FCD) in a time series prediction way. To detect the possible SOZ channels, the methodological framework of the proposed design will be more practical use in clinical applications.

As mention that, in the proposed design, we have used only one hour of interictal intracranial EEG data to identify the SOZ channel. In preprocessing step, we have used data segmentation and band pass filtering approach.

The experimental result shows that, the proposed machine learning based computer aided solution can identify the SOZ electrodes efficiently from short period of interictal iEEG recording, that provides epileptologists a great assistance and can increase the number of iEEG analysis for intractable epilepsy patients.

Declaration of Authorship

I, MOST. SHEULI AKTER, declare that the thesis I am submitting titled, "Machine Learning-Based Approach for Identification of Seizure Onset Zone in Patients with Focal Epilepsy from Interictal Intracranial EEG" is entirely my original work. Some parts of this thesis including methods, texts, figures, tables etc. has previously been submitted for a degree, which have been used at this thesis. I have clearly stated the contribution of other authors that I have included in my thesis.

Signed:

_____

Date:

_____

Acknowledgements

First and foremost, I would like to express the deepest appreciation to my respectable supervisor and mentor Professor Dr. Toshihisa Tanaka for his continuous support, motivation and direction. His guidance with patience, enthusiasm and immense knowledge is the ultimate outcome of this thesis. Without his generous help, it would have been impossible to finish this work.

I would also like special thanks to Assistant Professor Dr. Md. Rabiul Islam for his useful suggestions and directions to conduct the experiments during this journey. In addition, I would also like to thanks epilepsy surgeons, Dr. Hidenori Sugano, Dr. Takumi Mitsuhashi, and Dr. Yasushi Iimura for participating in interictal iEEG data recording and contributing to clinical supervision and evaluation.

I like thanks to Dr. Fabien Lotte (Inria, France), Dr. Anh Huy Phan (Skoltech, Russia) and Dr. Andrzej Cichocki (Skoltech, Russia) for their discussion and interesting research direction in this work. Also thanks to Kosuke Fukumori for his experimental and networking support. Besides, thanks to Dr. Wang Duo, Shintaro Ito, Dr. Khademul Islam Molla, and Kotaro Hayashi for their discussions in different times on this research.

I would wish to express deep gratitude to my family and friends for their mentally support. I like to special thanks all the teachers of my baby's daycare (Fureainomori daycare) for cooperation during this challenging journey.

I would like to many thanks the KDDI Foundation Scholarship, the Otsuka Toshimi Scholarship Foundation and JST CREST (JPMJCR1784) for partly supported financially during my PhD time. I would also thanks to Varendra University, Rajshahi Bangladesh for approving me of study leave to pursue this degree. Finally, thanks to all members of Biosignal Informatics Lab, Tokyo University of Agriculture and Technology. It would have been more difficult journey without the cooperation of you people.

Contents

List of Figures

List of Tables

Dedicated to my Parents and Family for their unconditional love and mental support...

Chapter 1

Introduction

1.1 Epilepsy

Epilepsy is one of the most common neurological disorder of the nervous system that creates severe effects to human brain [1], [2]. It is characterized as sudden and repeated seizure which may be caused by an excessive and synchronous electrical discharge of neurons inside the brain [3], [4]. A seizure is a sudden and uncontrolled electrical activity in the brain that produces disruptive symptoms. When a seizure occurs, epilepsy patients suffer from sudden and unexpected illness, during which they are unable to protect themselves and are vulnerable to suffocation, death, or injury due to fainting and traffic accidents [2], [5].

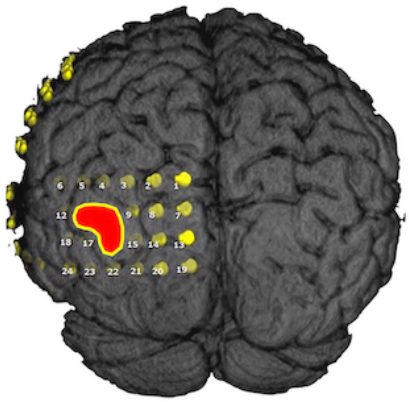

FIGURE 1.1: Brain image with epilepsy and its originated area. Red spot represent the SOZ area.

Seizures generally start in confined regions of the brain and may remain restricted to these areas or spread to other region of the brain. The area of brain from where seizures are initiated defined as seizure onset zone (SOZ) and it is usually localized by either scalp or invasive EEG [6]. The area of the cortex the (minimum amount of cortex) that must be resected (or completely disconnected) surgically to produce seizure

freedom is defined as epileptogenic zone (EZ). EZ is a dynamic concept that changes continuously over time [7]. According to the study of [8], [9] the most common epileptogenic region of the brain are the hippocampal formation and cerebral cortex. Fig. 1.1 shows the brain image with epilepsy seizure. The hotspot indicates the SOZ area from where seizure are originated.

Epilepsy is not a single disease, but a family of symptoms that share the feature of recurrent seizures. Epilepsy may be a result due to several factors, such as acquired structural brain lesions, inborn brain malformations, alterations in neuronal signaling, and defects in maturation and plasticity of neuronal networks [10]. Alternatively, it may also develop as a result of brain trauma such as a severe blow to the head, a stroke, central nervous system (CNS) infections, CNS malignancies, particularly cortically based tumors, such as gliomas and metastatic lesions, Alzheimer's disease etc.[11].

1.2 Social impact of epilepsy

The negative influence of uncontrolled (refractory seizure) seizures not only limited to the individual but also affect to their family members, friends, and the whole of society. The families and friends of people with refractory epilepsy experience life long concern and to ensure the safety of their loved one, they rearrange their lives. Several studies have revealed that, the major factors of quality of life such as job, family bonding, socializing etc. are on considerable risk in epilepsy patients [12], [13].

When seizure occurs, epilepsy patients suffer from sudden and unexpected illness. They become unable to protect themselves and vulnerable to suffocation, death or injury due to fainting and traffic accidents [2], [5].

Epilepsy can plays an vital rule in every year over other sudden illness related diseases. According to the study of [14]–[16], refractory epilepsy patients is associated with increased morbidity and mortality, serious psychosocial consequences, cognitive problems, and reduced quality.

Among the negative impact of epilepsy, people are frequently faced with adverse social difficulties. This type of difficulties are not only limited to adult people, but also impact in infancy. It have been reported that, children without epilepsy can show higher social competency compared to the patients with childhood epilepsy [17]. Because childhood epilepsy can impact seriously on the development of brain. It causes mental development disorder, severe motor and intellectual disabilities and cognitive dysfunction. In addition, the person who suffered with epilepsy in childhood, often faced many social difficulties after reaching the adulthood, even though his intellectual range is normal [18].

1.3 Diagnosis of epilepsy

For clinical diagnosis of epilepsy, epileptologist uses different techniques such as magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET), magnetoencephalogram (MEG), or electroencephalogram (EEG) [19]. Among them EEG signal plays an important role for detection and localization of epilepsy, because it gives temporal and spatial information about the brain as well as, measures differences in voltage changes between electrodes along the subject's scalp [4], [20], [21]. According to the information of Mayo clinic, the diagnosis process is define shortly in the following way:

- **Electroencephalogram (EEG):** To diagnose the epilepsy through the EEG test, electrodes are connected to brain by using the paste-like substance and cap. During this test, if someone have epilepsy then the

regular pattern of EEG waves have changes, even he is not having a seizure on that time.

- **Computerized tomography (CT) scan:** To get cross-sectional images of brain CT scan are uses. It can reveal the abnormalities of the brain such as tumors, bleeding and cysts etc. that might be causing of seizures.

- **Magnetic resonance imaging (MRI):** To create a detailed view of brain through powerful magnets and radio waves doctor uses MRI images. To detect the lesions or abnormalities of brain, doctor uses these MRI images.

- **Positron emission tomography (PET):** To detect the abnormalities and visualize the active areas of the brain PET scans use a small amount of low-dose radioactive material. For PET scans, these radioactive material are injected into a vein.

1.4 Epilepsy treatment

According to World Health Organization (WHO), it is the second most common neurological disorder behind stroke. Roughly 50 million people all over the world have been diagnosed with epilepsy [1], [10]. As in studies [22], approximately one in every one hundred individuals suffers from epilepsy.

Increased mortality is associated with epilepsy patients. Due to the unpredictability of seizures, the primary burden of the disease is reduced quality of life and lost productivity. Deaths due to to epilepsy are blamed on seizure-related accidents, status epilepticus (very long seizures that become life threatening). and the syndrome known as 'SUDEP' (sudden unexpected death in epilepsy) that can be associated with cardiorespiratory dysfunction [23].

There are a different categories of epilepsy treatment that are shortly describe in these section.

1.4.1 Medication

There are several options to treat the epilepsy patients, medication can be effective in eliminating or significantly reducing seizures. According to the information of Mayo clinic, by taking one anti-epileptic medication most people can become seizure-free. But, for many patients after a few months, the utility of AEDs (medication) gradually diminis. Sometimes utility of medication shows some side effects (fatigue, dizziness, weight gain, loss of bone density, skin rashes, loss of coordination, speech problems, memory and thinking problems etc (information based on Mayo clinic)).

According to the studies in [24] showed that, one-third of adults with epilepsy have not adequate control of seizures with medication. Recently, a study published in 2018 showed that despite availability of many new antiepileptic drugs with varing mechanisms of function, but overall outcomes in newly diagnosed epilepsy have not improved. [25]. The seminal work of [26] reported that 43 out of 143 case studies (30%) achieved no control of their seizures despite the recent discovery of bromide as a promising new AED. Therefore, the effectiveness of medication in reducing seizure rates has remained unchanged for over one hundred years.

1.4.2 Surgery

When medication fails to control the epilepsy, physician chooses the surgical treatment as a possible solution to recover from epilepsy [3], [4] . Surgical resection of epileptic parts of the brain may be considered if the seizure focus is able to be located and removed safely. Doctors takes decision for surgery (based on information of Mayo clinic):

- When seizure are originated from a small, well define area of brain.
- When brain does not response for some vital functions such as as speech, language, motor function, vision or hearing.

However, for some patients with epilepsy are not suitable surgical candidates (eg. in cases with generalised epilepsy or where the focus is in an inoperable region of the brain), and surgery does not always provide complete seizure freedom [27]. In another studies it has been showed that successful surgery tends to relieve depression and anxiety[28] and 70% of people who have surgery become sezure free.

1.5 Necessity of machine learning based epilepsy detection framework

For surgical treatment of epilepsy, the localization of SOZ zone or epileptogenic zone is very important step. Usually, epileptologist localized these zone by analyzing iEEG data. It has been said that, the specialist of epilepsy (epileptologist) is the translator of iEEG signal. For correctly interpretation of iEEG signal, the epileptologist need much training for many years. In addition, iEEG are recorded from 40∼148 electrodes positions for few weeks, at least 2∼3 days. So, to localization of the SOZ zone manually, the epileptologist requires visual observation on long-term multichannel iEEG data.

Moreover, there is a lacking of clinical expert for such diagnosis. According to the information of Japan Epilepsy Society in 2019, there are only 689 specialist all over the Japan. In some prefacture in Japan, their are only one specialist. So localization or detection of SOZ or epileptogenic zone by visual observation of long time iEEG data is consuming and laborious process [4], [29], [30] and it puts heavy burden on specialist of epilepsy and reduces their efficiency. Therefore, a computer-aided solution with an effective algorithm

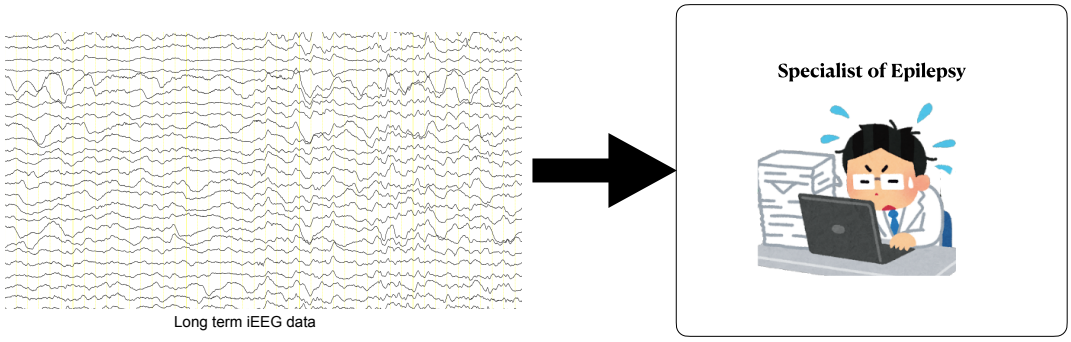

FIGURE 1.2: Process to interprete the iEEG data.

that uses the iEEG signal to localize the epileptic focus would be invaluable. Fig. 1.2 shows conventional way for interpretation of long term iEEG data by the epileptologist

1.6 Epilepsy seizure related existing studies

1.6.1 Epilepsy seizure detection studies

Considering the low frequency components (LFC)-related epilepsy seizure detection framework, several studies have been proposed [31]–[33]. Nicolaou et al. proposed the method to detect the epileptic vs non-epileptic (normal subjects) activities using Bonn dataset [34]. The datasets consists of five subjects with single-channel surface EEG at a sampling rate of 173.61 Hz denoted as Set A (healthy awake and eyes open), Set B (healthy awake and eyes closed), Set C (epileptic seizure-free interictal), Set D (epileptic seizure-free interictal), and Set E (epileptic seizure activity). A two class problem was formed by assigning the class labels to epileptic activity (Set E) against the activity from other subjects (A, B, C, and D). Therefore, 4-different combinations such as E vs A, E vs B, E vs C, E vs D, were performed to detect epileptic vs non-epileptic activities. Guo et al. proposed MLPNN classifier with multiwavelet entropy to detect normal vs ictal [35] with similar dataset, improving the the accuracy of 99.06%. Hamad et al. proposed discrete wavelet

transform (DWT) based features with radial basis function SVM for recognition of epilepsy [36]. Sharathappriyaa et al. used auto-encoders based automatic seizure detection system using harmonic wavelet packet transform and fractal dimension based features and obtained 98.67% accuracy [37]. A CNN based machine learning method was implemented in [4] for epilepsy detection with significant classification accuracy. Ullah et al. established deep learning based automated epilepsy detection system using Bonn dataset and achieved the accuracy of 99.1% [38]. Recently, Li et al. combined three types of faster feature-extraction with EMD decomposition for two datasets (Bonn and Qilu Hospital) to measure the dynamics nature of EEG signals for epileptic and non-epileptic patients [31]. The MLPNN-based classifier was used to detect normal vs ictal patterns in EEG. The same feature-extraction methods with DWT and FNN were proposed to improve the usability and performance of the system in the epilepsy study by Hasan et al. [32]. However, the main target of the above epilepsy studies using Bonn dataset was to detect the epilepsy patients (normal vs epilepsy subject) rather than the identifying of SOZ, which limits the clinical applications.

1.6.2 Low frequency component related focal seizure (or SOZ) detection studies

For identification of epilepsy focal also referred to SOZ, several studies have been proposed by considering the low frequency components (LFC) with intracranial EEG signal [39]–[42]. Bhattacharyya et al. [43] have proposed a machine learning based focal EEG identification system using least-squares support vector machine (LS-SVM) classifier. They have used Bern-Barcelona datasets which consists of 7500 pairs of focal and non-focal iEEG signal. For 50 pairs of focal and non-focal EEG signals, their achieved classification accuracy of 90%, sensitivity and specificity of 88% and 92%, respectively. The

same dataset were used by Arunkumar et al. for clssification of focal and non-focal iEEG signals [41]. In their proposed design, they used entropy based feature extraction methods with six different classifiers including naïve bayes (NBC), radial basis function (RBF), support vector machine (SVM), KNN classifier, non-nested generalized exemplars classifier and best first decision tree (BFDT) classifier. They achieved highest accuracy of 98%, sensitivity of 100% and specificity of 96% with non-nested generalized exemplars classifier (NNge). Sharma et al. [39] have introduced entropy based feature extraction method with SVM classifier for automatic identification of focal and non-focal pattern of signal. They have also used Bern-Bercelona dataset with 87% of accuracy. However, the Bern–Barcelona dataset be formed of nearly 20-s of iEEGs with a pair (two-channel iEEG) of focal and non-focal channels and upper and lower cut off frequency is between 0.5 and 150 Hz. The bi-variate focal and non-focal channels were recorded at the epileptic and non-epileptic zone from the brain. Recently, EMD (emperical mode decomposition) and bivariate EMD a time domain multiband decomposition method have been proposed in studies [40], [44] to detect the epilepsy seizure. For feature extraction, they used different entropy based feature extraction methods and improve the accuracy of the system. Although all of the above studies achieved significant classification, but the limitations is that, they have used low frequency bandwidth between (0.5 to 150 Hz) interictal iEEG data, which limit the clinical use. Moreover, they have evaluate their system with well balanced data which are contrary to most real world problem.

1.6.3 High frequency oscillations (HFOs) related detection

However, the use of a low frequency band that partially excludes higher frequency components has some restrictions in real world applications. Recently, several developed methods on epilepsy detection [45]–[47] have illustrated that, the high frequency oscillations (HFOs) including ripple (100–250 Hz) and fast ripple (250–600Hz) bands carries crucial biomarker to detect the seizure onset zone (SOZ) as well as to guide epilepsy surgery. There was more evident that, the fast ripple band contained more repetitive waveform pattern compared to ripples band and it complicates the clinical use of HFOs as valid biomarkers [48]–[50].

Most recently, Zuo et al. [51] have proposed the CNN-based method for identifying ripple and fast-ripple band and compared their studies to other four proposed HFOs detection methods in the RIPPLELAB toolbox [52]. Most of the above HFOs related studies calculate the baseline from long term iEEG data and detect automatically ripple and fast ripple separately. However, recent finding of HFOs related studies limits the existing systems for clinical use.

1.7 Our contributions

1.7.1 High frequency component related SOZ detection with entropy based features

Through motivated by recent finding of HFOs related studies [45], [47], in this study, we have proposed high frequency component (HFC) based SOZ identification method. It is known that, the activity in high frequency components, including ripple and fast ripple bands, of interictal iEEG, are associated to the epileptic seizures. In our proposed methodological framework,

we have used entropy based feature extraction methods with filter bank approach. Based on different epilepsy related studies [39]–[41], [44], we hypothesized that, the combination of entropy features with filter-bank approach are effective for identifying the epileptic event. To select the prominent entropy features, we have used sparse linear discriminant analysis (sLDA). Eight patients with intractable epilepsy are used to evaluate the method. To evaluate the proposed design, we have used 10-fold cross validation with SVM classifier. As mentioned that, the dataset for all patients are highly imbalanced, the number of non-SOZ electrodes are much higher compared to the SOZ electrodes. To solve this imbalanced problem, we have used adaptive synthetic oversampling approach (ADAYSN) in the training stage of SVM classifier. The detail description of this study with result and graphical representation shown in Chapter 4. Considering the noise-robust features and the reduction of method complexity, the multi-band feature-extraction method has a great potential as the basis for designing a computer-aided solution for localizing SOZ electrodes.

1.7.2 High frequency component related SOZ identification with efficiently work statistical features

Our previous study [53], which are discussed in Sec. 1.7.1, suggests that entropy features in HFCs are still effective in identifying SOZ electrodes. The study [53] proposed eight entropy-based feature extraction methods for identification of SOZ electrodes. However, that framework has some shortcomings, among them the significant one, the parameters selection problem. Basically, the performance of entropy estimation sharply depends on the parameters selection [54], [55]. Another shortcomings is the higher computational cost. To utilize the high-frequency with a range from 100 to 600 Hz in each segment of multi-channel iEEG, the number of sample points in each

segment are large enough, which lead to the higher computational cost [53]. Furthermore, the identification of SOZ channels are acquired based on the number of identified segments on SOZ and non-SOZ channels, and the detected decision is make based on an SVM classifier based hard thresholding.

To address the problems described in above, we have hypothesized that simple statistical features are useful in SOZ identification due to these features have already been applied in other context of epilepsy studies [31]–[33], [56], [57]. Thus, the contributions of our second proposed design are the following:

- We have proposed twelve feature extraction methods, including of nine statistical and three information theoretic entropy features that are significant to characterize the epileptic signals.
- A data-driven grid-search method using mutual information (MI) scores has proposed to optimize prominent bands and features jointly that are still not reported in previous SOZ detection studies.
- We have used standard state-of-the-art LightGBM classifiers to identify the possible SOZ electrodes.

1.8 Organization of the thesis

As depicted in Fig. 1.3, this thesis is organized as follows: In Chapter 1, we have discussed about the background of this work, its problem and possible solutions. In Chapter 2, we have introduced the basics of epilepsy and its categorization. In Chapter 3, we have reviewed the recent filter bank techniques with the other related methods. Multiband entropy based feature extraction methods for identification of SOZ electrodes have discussed in Chapter 4. In

Chapter 5, we have discussed about the statistical features based SOZ identification methodological framework with selection of prominent features and band. In Chapter 6, we have concluded the thesis.

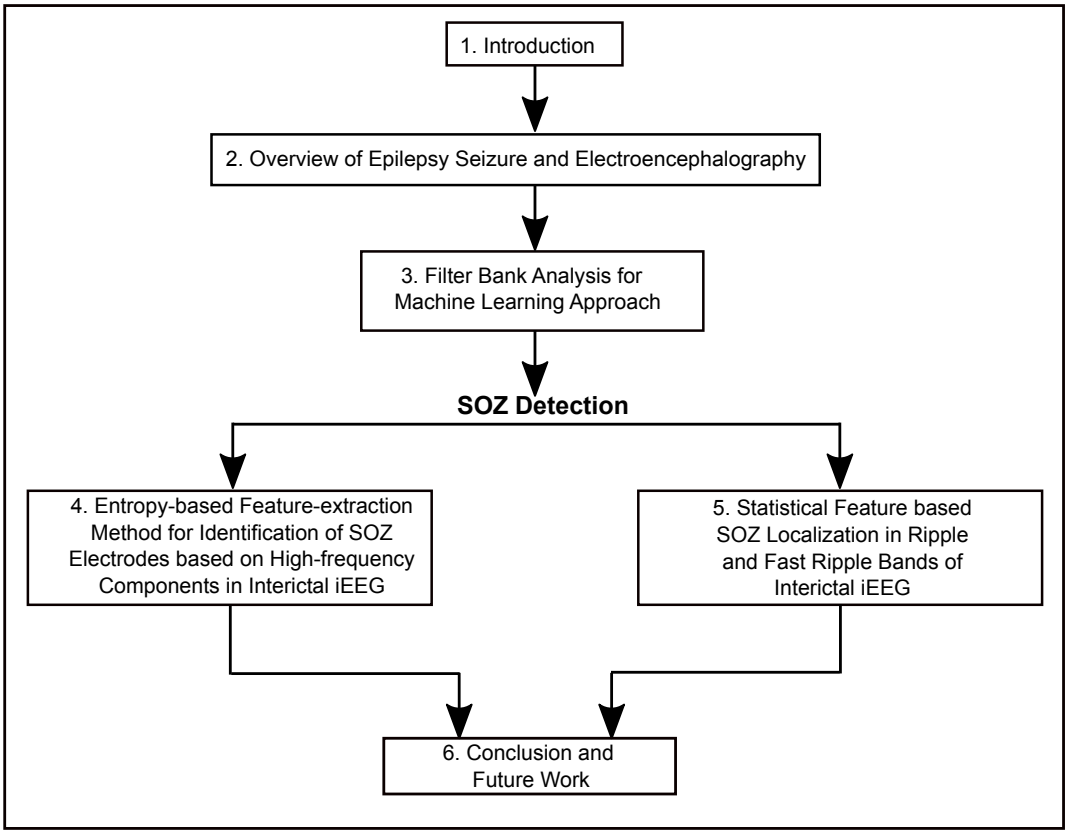

FIGURE 1.3: Diagram for the organization of the thesis.

Chapter 2

Overview of Epilepsy Seizure and Electroencephalography

In this chapter, we have discussed about the epilepsy seizure originated area such as seizure onset zone (SOZ), epileptogenic zone (EGZ) in Sec. 2.1. Sec. 2.2 presented the epilepsy seizure categorization based on international league against epilepsy (ILAE). In Sec. 2.3, discussed about the categorization of EEG including scalp electroencephalogram (EEG) and intracranial electroencephalogram (iEEG).

2.1 Different cortical area related to epilepsy seizure

2.1.1 The irritative zone

The irritative zone is defined as the area of cortex that creates interictal electrographic spikes and it is generally called mini-seizures [58]. These zone usually localized by scalp or intracranial electroencephalogram (EEG or iEEG), magnetoencephalography (MEG) [59] or spike-triggered functional MRI (fMRI) [60]. If irrititive zone are created in an eloquent area of cortex and these are of enough 'strength', spikes can give raise to clinical symptoms. For example, patients with mesial temporal lobe epilepsy, exhibit temporal epileptiform

discharges. On the otherhand, isolated or independent spikes will not produce any clinical symptoms if they located in silent or eloquent cortical area.

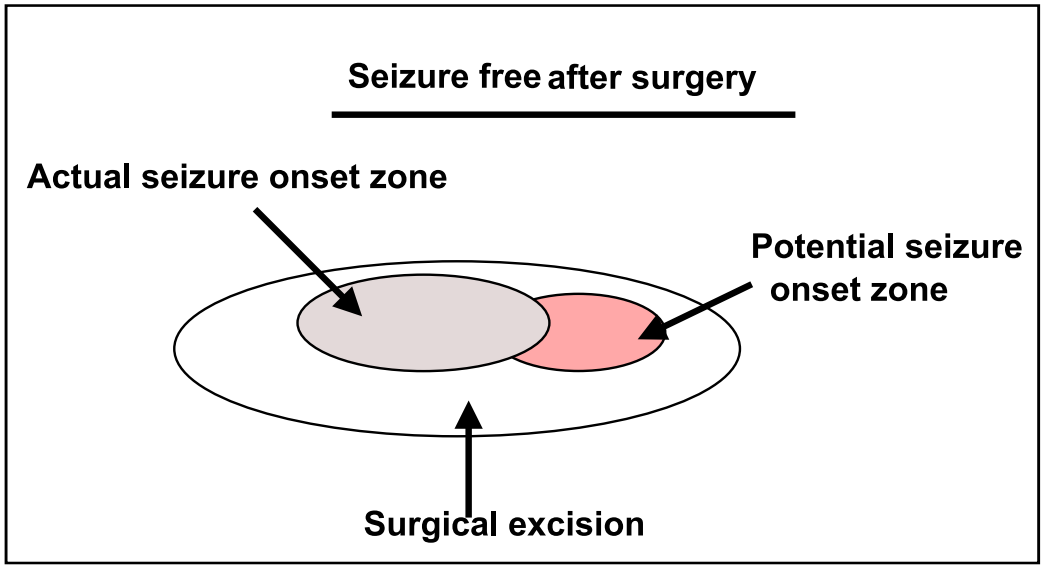

FIGURE 2.1: The example of SOZ with surgical excision [61]

2.1.2 The seizure onset zone

The seizure onset zone represent the cortical area of brain from where clinical seizures are originates [58]. The seizure onset zone is measured by EEG (scalp or invasive) and location of these zone can be determinated with ictal SPECT (single photon emission computed tomography). Usually, it is the part of the irritative zone and it produce repetitive spikes that have sufficient strength to generate clinical ictal symptoms when attacking eloquent cortex. Invasive electrodes are more sensitive compared to scalp electrodes for detecting the SOZ zone. Talairach et al. [62] reported to their studies that the seizure-onset zone would be a reliable index of the location and extent of the epileptogenic zone. In their study, they defined the seizure onset zone as epileptogenic zone. Lüders et al. [61] enriched this notion by identifying five cortical zone such as irritative zone (area of cortex which generates interictal spikes), seizure onset zone (area of cortex that initiates clinical seizures),

symptomatogenic zone, lesion, and functional deficit zone. Lüders et al. observed that, for some cases the complete surgical removal of seizure onset zone (SOZ) does not lead to seizure freedom, because there are some other areas in the brain that were closest to the removed SOZ zone, are triggering the epileptic seizures [61]. They arrived to the conclusion from their study that SOZ and epileptogenic zone are not same concepts and they separated the SOZ into actual and potential SOZ. They suggested that the complete removal of both the actual SOZ and potential SOZ through surgery may result in seizure free (see in Fig. 2.1)

2.1.3 The epileptogenic zone

The epileptogenic zone (EGZ) is the cortical area of the brain that is inevatable for the generation of epileptic seizures. It may comprise an actual epileptogenic zone and potential epileptogenic zone. The actual epileptogenic zone is the area of cortex that generating seizures prior surgery, whereas potential epileptogenic zone refers to the area of cortex that may produce seizures after the presurgical SOZ has been resected. To understand the seizure mechanisms and for complete abolition of seizures through surgical removal of seizure focus, the localization of the boundary of EGZ is needed [61]–[63]. Since, there is no directly diagnostic modality currently available for measuring the entire EGZ, we need to infer its location indirectly by defining the other zones such as potential SOZ, actual SOZ or irrititave zone. Therefore, epilepsy surgeon and epileptologists need to define the location and the extension of EGZ by observing the SOZ before the epileptic focus resection for patients with medically intractable epilepsy. At the end, to what extent area in the brain should be considered as EGZ and resected for good surgical outcome is still an open research.

2.2 Epilepsy seizure and its categorization

Epilepsy is the disease related with spontaneously recurring seizures in which activity of brain becomes abnormal. Any ages people both males and females can be affected with epilepsy. (According to the definition of Miyo clinic)

In 2017, the international league against epilepsy (ILAE) published in two articles an updated classification of seizures and the epilepsies, together with an instructional manual on how to apply the seizure classification [64]–[67].

These seizure types are describe briefly in the following way:

- **Focal onset seizures:** Focal seizures can originate from one cerebral hemisphere or either in a specific place in the brain [64]. It can be divided into two groups: Focal onset aware seizures, focal onset impaired awareness.
  - **Focal onset aware seizures:** When a person is awake during the seizure, then it's called a focal aware seizure. In this type of seizure, person don't loss their consciousness and this type of seizure also call simple partial seizures.
  - **Focal onset impaired awareness:** When person loss their consciousness or become confused then this type of focal seizure is called focal impaired awareness seizure. Previously, it was called complex partial seizure.
- **Generalized onset seizures:** In these type of seizure, a group of cells on both side of the brain are affected at the same time. These type of seizure included tonic-clonic, absence or atonic ete.
- **Unknown onset seizures:** Unknown onset seizures are such type of seizure when its originating location is not known or when the physician has not yet congeal adequate clinical information to be certain about the epilepsy classification [68].

The above three onset types seizure are further classified as motor onset seizure and non-motor onset seizure. Different types of seizures are responsible for different types of motor or non-motor behaviour. For focal seizure, a particular region of the brain are affected. The affected area of the brain may be occipital lobe, temporal lobe, frontal lobe or parietal lobe and based on the affected area of the brain, the behaviour of the seizure are changed. For example, automatism changes may be occur due to the temporal lobe seizures whereas frontal lobe seizures often lead to hyperkinetic seizures.

- Myoclonic seizures: A sudden, short contractions of muscles that lasting ≤ 0.5 s [69]. It is a twitch-like contraction and in most cases awarness is not impaired. Myoclonic seizures typically includes limbs and shoulders and can potentially occur in clusters.

- Tonic: An increase in one or multiple muscles contraction often resulting in a change of posture. This type of seizures lasting from a few seconds to some minutes [64] and mostly awareness is not loss.

- Clonic: It is a series of myoclonic contractions where symmetrical or asymmetrical jerking occur of the same group of muscles. In clonic seizures, typically the whole body are affected involved and generally awarness is impaired.

- Tonic-clonic: These type of seizures are the combination of a tonic and a clonic seizure where muscles first stiffen (tonic phase), and then start to jerk (clonic phase) [65]. In tonic-clonic seizures, patients typically lose consciousness.

- Hyperkinetic: Seizures with agitated thrashing and repetitive movements. These type of seizures also involves in leg pedaling movements.

- Automatisms: Automatisms seizures are frequently originating from the temporal lobe of the brain and it impaired by performing one or

more short unconscious behaviors. Smacking, swallowing, chewing and rubbing fingers/hands etc. are the symptoms these type of seizures.

2.3 Categorization of electroencephalogram (EEG)

Basically two types of EEG are used in EEG based diagnosis system. 1). scalp EEG (EEG) and ii). intranial EEG (iEEG).

2.3.1 Scalp electroencephalogram (EEG)

The non-invasive measure of electrical activity of the brain is represented by scalp EEG. It is generally measured by those electrodes that are symmetrically arrayed on the scalp. EEG measures voltage fluctuations resulting from ionic current within the neurons of the brain [70]. Clinically, EEG refers to the recording of the brain's spontaneous electrical activity over a period of time and recording is done from multiple electrodes placed on the scalp [70].

For scalp EEG signal, there are some limitations for both origin and characteristics of neural activity. In particular, in scalp EEG signal, the activity of neurons that are closest to scalp surface are more visible compared to the neurons activity buried within deep brain structures. Besides, the cerebrospinal fluid and skull surrounding the brain act as attenuators that comprehensively reduce the amplitude of high frequency neural oscillations. As a result, some types of epilepsy seizures that occur within the deep region of brain can not be observed by scalp EEG.

2.3.2 Intracranial electroencephalogram (iEEG)

In iEEG signal, electrodes are placed on the cortex of brain or deep within the brain structure and like scalp EEG, it also gives a spatial and temporal information about the electrical activity of neurons. It is less affected by physiologic and environmental artifacts compared to scalp EEG. Fig. 2.2 shows the recording of iEEG data, where electrodes are directly placed the on the cortex of brain.

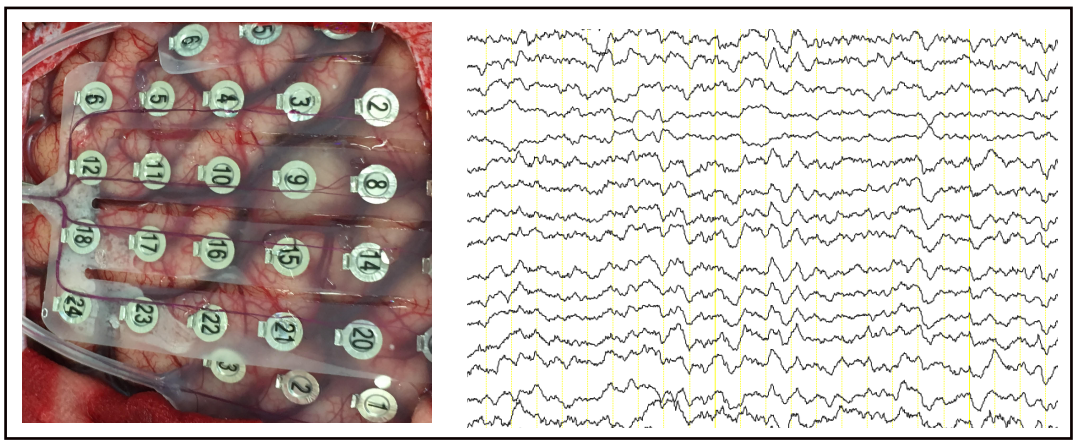

FIGURE 2.2: Example of intracranial EEG data recording [Juntendo University Hospital]

iEEG are highly invasive and since electrodes are placed in a limited region of brain sites at any given time, so it provides higher spatial resolution and worse spatial coverage compared to scalp EEG.

In iEEG, the neuronal hypersynchrony associated with a seizure is noticed in lower time (tens of seconds) than the scalp EEG [4]. Moreover, higher spatial resolution of iEEG signal allows the recording of such abnormal, non-seizure activity that are not visible within the scalp EEG [71], [72].

Chapter 3

Filter Bank Analysis for Machine Learning Approach

In this chapter, we have introduced filterbank analysis and some other related methods, which can be used to the real world signals.

3.1 Multi-band analysis

Filter-banks are used for spectral decomposition and composition of signals, that are the arrangements of lowpass, bandpass, and highpass filters [73]. In many mordern signal processing applications such as mobile communication, speech signal processing, neuroengineering, audio and image coding, they plays a vital role.

Filter bank can be uniform and non-uniform. In case of uniform filter banks, the bandwidth and sampling rate of all filters are same. From the point of view of implementation, uniform filterbank are preferred frequently. On the otherhand, octave-spaced or wavelet filter bank is one of the popular example of non-uniform filter bank. In this thesis, to design a framework, we have used Butterworth filter.

3.2 Finite impulse response (FIR) filter

FIR filter is a primary type of digital filter whose impulse response remain in finite duration and it used in signal processing. Suppose that, $x[l]$ is the input signal and $y[l]$ is the output signal of FIR filter. Then the output sequence of each value for FIR filter with order L is computed in the following way:

$$y[l] = w_0 x[l] + w_1 x[l-1] + \ldots + wLNx[l-L] = \sum_{j=0}^{L} w_j.x[l-j] \tag{3.1}$$

where w_j represents the coefficient of j^{th} instant.

3.3 Infinite impulse response (IIR) filter

IIR filters are usually used when computational resources are at a premium [74]. Since, the stable, causal IIR filters cannot have perfectly linear phase, so IIR filters tend to be avoided when linearity of phase is a necessity [74]. These filters have the feedback and are acquinted as recursive digital filters. For the same IIR filters have much better frequency response compared to the FIR filters due to the the recursive part of filter. For the IIR filter, let us consider, $x[l]$ and $y[l]$ are a input signal and output signal. So the output signal of IIR filter $y[l]$ can be define as:

$$y[l] = \frac{1}{v_0}(w_0 x[l] + w_1 x[l-1] + \ldots + w_R x[l-R] - v_1 y[l-1] - v_2 y[l-2] - \ldots - v_S y[l-S]) \tag{3.2}$$

where R and S indicates the feedforword and feedback filter order. The w_i and v_i are the feedforward and feedback filter coefficients.The eq. (3.2) can

also be expressed as:

$$y[l] = \frac{1}{v_0}(\sum_{j=0}^{R} w_j x[l-j] - \sum_{k=1}^{S} v_k y[l-k]) \tag{3.3}$$

where w_j and v_k are the feedforward and feedback filter coefficients. The eq. (3.3) can also be compressed as:

$$\sum_{k=0}^{S} v_k y[l-k] = \sum_{j=0}^{R} w_j x[l-j] \tag{3.4}$$

To find the transfer function of the filter, taking the Z-transform of each side of eq. (3.4) and expressed as:

$$\sum_{k=0}^{S} v_k z^{-k} Y(z) = \sum_{j=0}^{R} w_j z^{-j} X(z) \tag{3.5}$$

So, the transfer function can be define as:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{j=0}^{R} w_j z^{-j}}{\sum_{k=0}^{S} v_k z^{-k}} \tag{3.6}$$

In most IIR filter designs, the coefficient v_0 is consider as 1. So the more traditional design of IIR filter transfer function are:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{j=0}^{R} w_j z^{-j}}{1 + \sum_{k=1}^{S} v_k z^{-k}} \tag{3.7}$$

In this thesis, we have used IIR butterworth band pass filter to design the framework.

3.4 Discrete wavelet transformation

Discrete wavelet transforms (DWT) are widely used in many epilepsy related study [75]–[77]. The advantage of DWT over fourier transform is that, it have

a good time-frequency localization, multirate filtering and scale space analysis [76]. It has the similarity with Fourier transform, but the difference is that transform decomposes the signal into sines and cosines function localized in Fourier space. Moreover, there are the similarity between the wavelet transform and Fourier transform, but the main difference is that Fourier transform decomposes the signal into sines and cosines function localized in Fourier space. On the other hand, the wavelet transform uses functions that are localized in both the real and Fourier space [78].

Since window size of WT are variarble, so it provides more flexible way of time-frequency representation of a signal compared to STFT (short term Fourier transform). In wavelet transform, long and short time windows are used to get a low and high frequency information respectively. Thus, it gives the important frequency and time information for low and high frequencies and this charactristics makes the WT more appropriate for analysis irregular data pattern [75].

The discrete wavelet transform (DWT) of a signal $x\left[l\right]$ is written as:

$$Y(a,b) = \sum_{l \in z} x\left[l\right] \Psi_{a,b}\left[l\right] \tag{3.8}$$

Where a represent the dilation or scale, b indicates the translation and

$$\Psi_{a,b} = \left(\frac{1}{\sqrt{a}}\right) \Psi^{*}\left(\frac{l-b}{a}\right) \tag{3.9}$$

DWT decomposes a signal into approximate and detail coefficient.

DWT uses high pass filter to analyze high frequency content and low pass filter for low frequency contents. Signal resolution is changed by filtering operation. And up sampling and down sampling operations are performed on signal to change a scale.

3.5 Empirical mode decomposition

In recent years, empirical mode decomposition (EMD) is gained popularity in iEEG signal related epilepsy study [79]. EMD is a data dependent decomposition method that decompose a signal into a set of subband signal and each subband signal is called the intrinsic mode functions (IMFs). The IMFs are amplitude and frequency modulated components [80]. EMD is suitable for analysis of univariate nonstationary and nonlinear signal. [80]. But, for bivariate or multivariate signal the frequency band of each IMF is not consistant due to applying it independently of each channel of signal. To overcome the problem, Rilling et al. proposed an extension of EMD to bivariate time series, namely bivariate empirical mode decomposition (BEMD) that generalizes the rationale underlying EMD to the bivariate framework [81]. BEMD has been applied in many application such as image segmentation [82], image fusion[83], image watermarking [84], turbine condition monitoring [85], grinding chatter detection [86] etc. It also have been applied in epilepsy related study [40].

3.5.1 Bivariate empirical mode decomposition (BEMD)

The BEMD decomposes a multiple signal into a set of band limited intrinsic mode function (IMFs). For each IMFs, the following two conditions must satisfies [86]:

* The number of extrema and the number of zero crossings must be equal, or differ at most by one.

* The mean value of the envelope at any point defined by the local maxima and the local minima must be zero.

In addition, BEMD technique need to project a multiple signal on a set of directions and then requires to apply shifting process to the projected components. For a given bivariate signal, $s(l) = x(l) + iy(l)$, the BEMD process can be describe as follows [86] :

1. A set of projection directions ψ_u:

$$\psi_u = 2u\pi/L, 1 \leq u \leq L \tag{3.10}$$

2. Project the complex signal $s(l)$ on directions ψ_u:

$$p_{\psi_u}(l) = Re[e^{-i\psi_u} s(l)], \tag{3.11}$$

3. obtain all partial maxima of $p_{\psi_u}(l) : (l_j^u, p_j^u)$, where j indicates the number of individual partial maximum points.

4. By cubic spline interpolation, interpolate the set of points $(l_j^u, e^{-i\psi_u} p_j^u)$ to obtain the partial envelope curve $e^{-i\psi_u}(l)$ in direction ψ_u.

5. Steps 2-4 repeat until the envelop curves in all L projections are attained.

6. For all envelops curves, calculate the average:

$$\bar{u}(l) = \frac{1}{L} \sum_{r=1}^{L} e_r(l) \tag{3.12}$$

7. To obtain $f(l)$, subtract the mean $\bar{u}(l)$ from $s(l)$:

$$f(l) = s(l) - \bar{u}(l) \tag{3.13}$$

8. Test whether the $f(l)$ is an IMF or not. If $f(l)$ is not IMF, then replace $s(l)$ with $f(l)$ and repeat the Step from 2 to 7 until $f(l)$ is an IMF. When $f(l)$ satisfies the IMF condition then:

$$c_1(l) = f(l) \tag{3.14}$$

and residual signal $r_1(l)$ as:

$$r_1(l) = s(l) - c_1(l). \tag{3.15}$$

9. To find the second IMF $c_2(l)$, consider $r_1(l)$ as the original signal and repeat the above procedure until it satisfies IMF condition. After getting the second IMF, the residual component $r_2(l) = r_1(l) - c_2(l)$.

10. Repeat all the previous steps until all IMFs are obtained.

For a signal $s(l)$, the BEMD can be expressed by the procedure:

$$s(l) = \sum_{k=1}^{n} h_k(l) + r_n(l) \tag{3.16}$$

where $h_k(l)$denotes the kth IMF and $r_n(l)$ denotes a non-zero mean low-degree polynomial residue.

3.6 Limitation of methods to design ML based multi-band analysis

The wavelet and EMD based multivariate approach are appropriate for non-linear and non stationary EEG signal, because they are fully data adaptive and extract more useful information across space, time and frequency. But, all of these methods have some limitations for machine learning approach:

1. For both train and test set, these methods do not insure the same bands of the components.

2. The computational cost is high compared to the filter bank approach.

3. These methods increase the system complexity.

4. For real time implementation, multiband filter bank approach are more appropriate compared to other methods.

Considering the advantage of filter bank approach over wavelet, EMD, BEMD and MEMD, this thesis has proposed the filter-bank approach to detect the SOZ electrodes through the machine learning methods.

3.6.1 Linear discriminant analysis

Linear discriminant analysis (LDA) [87]–[89] is a supervised subspace learning method used for dimensionality reduction and classification. The aim of this methods is to maximize the between-class scatter matrix and within-class scatter matrix. It find the linear transformation $\upsilon \in \mathbb{R}^{d\times l}$ and maps $\boldsymbol{y}_j$ in the d-dimensional space to a l-dimensional space. Mathematically, it can be expressed as [90]:

$$\arg\max_{\mathbf{w}} tr\left(\left(\upsilon^T \mathbf{S}_w \upsilon\right)^{-1}\left(\upsilon^T \mathbf{S}_b \upsilon\right)\right) \tag{3.17}$$

where $\mathbf{S}_b$ represent the between-class scatter matrix and $\mathbf{S}_w$ indicates the and within-class scatter matrix. The $\mathbf{S}_w$ and $\mathbf{S}_b$ can be defined as:

$$\mathbf{S}_b = \sum_{k=1}^{c} n_k \left(\mu_k - \mu\right)\left(\left(\mu_k - \mu\right)^T, \tag{3.18}$$

$$\mathbf{S}_w = \sum_{k=1}^{c}\sum_{j\in C_k} \left(\mathbf{y}_j - \mu_k\right)\left(\left(\mathbf{y}_j - \mu_k\right)^T, \tag{3.19}$$

where C_k is the k-th class index, μ_k is the mean vector and n_k represent the size of k-th class. in the input data space, $\mathbf{Y}$. By using eq. (3.18) and eq. (3.19) we get:

$$\arg\max_{\mathbf{w}} tr\left(\left(\upsilon^T \mathbf{S}_t \upsilon\right)^{-1}\left(\upsilon^T \mathbf{S}_b \upsilon\right)\right) \tag{3.20}$$

where $\mathbf{S}_t$ represent the total scatter matrix which is define as:

$$\mathbf{S}_t = \sum_{j=1}^{n} \left(\mathbf{y}_j - \mu\right)\left(\mathbf{y}_j - \mu\right)^T \tag{3.21}$$

Note that, $\mathbf{S}_t = \mathbf{S}_w + \mathbf{S}_b$.

Chapter 4

Entropy-based Feature-extraction Method for Identification of SOZ Electrodes based on High-frequency Components in Interictal iEEG

In this chapter, we have discussed about the detection of seizure onset zone (SOZ) using multi-band entropy-based features with machine learning in high frequency components of interictal clinical iEEG. Considering the noise-robust features and the diminishing of design complexity, the multi-band feature extraction method has a great possibility as the basis for designing a computer aided method for identifying SOZ electrodes. In this chapter, we have discussed about the proposed design architecture in Sec. 4.1. It includes the dataset, focal cortical dysplasia (FCD), multiband analysis and entropy based feature extraction methods. We have discussed about the the feature selection and imbalanced learning problem in Sec. 4.2 and 4.3 respectively. Evaluation part is introduced in Sec. 4.4. Finally experimental result and discussion are presented in Sec. 4.5 and Sec. 4.6.

4.1 Proposed design architecture

In this section, we have discussed about the architecture of our proposed design shown in Fig. 4.1. The multi-channel interictal iEEG data are recorded for at least three days until an adequate number of habitual seizures were obtained for analyzing. Among them, to design the SOZ detection framework, we have used only 30-minutes of interictal iEEG signal. As a preprocessing step, we have splitted the 30-mins interictal iEEG signal into 20-s segments resulting in total 90 segments. Then a third-order Butterworth bandpass fitter are applied to extract the high-frequency components (100–600 Hz) from each interictal iEEG segment. Considering the design performance as well as the reduction of the complexity, we have divided the high-frequency bands, including ripple (100–250 Hz) and fast ripple (250–600 Hz), into 10 subbands, each of which has a band width of 50 Hz. The subbands are labeled as $S_1, S_2, \ldots, S_L$, where L is the total number of subbands ($L = 10$). After that, we have used the eight entropy based feature extraction methods to extract features from each subbands. After extracting features from each subbands, feature selection method such as sparse linear discriminant analysis is applied to select the prominent features. Finally, we have concatenated the subbands and selected features for classifying SOZ and non-SOZ channels using SVM classifier based hard thresholding.

4.1.1 Dataset

In this study, eight patients data are collected from Juntendo University Hospital in Tokyo, Japan that was approved by the ethics committee of Juntendo University Hospital as well as the Tokyo University of Agriculture and Technology, Tokyo Japan. All process are accomplished in accordance with relevant guidelines and act. All the patients are signed the informed consent for a research protocol. During the pre-surgical evaluation, several non-invasive

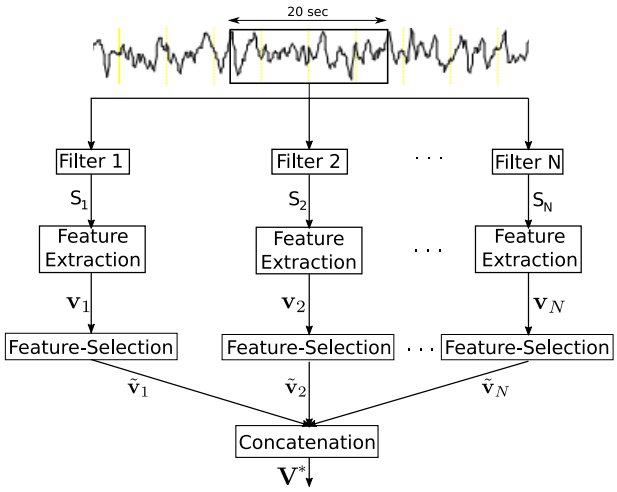

FIGURE 4.1: The different components of the proposed design for SOZ electrodes identification.

pathology protocols, for example seizure semiological evaluation, interictal scalp EEG, MRI, molecular imaging, and psychomotor-development testing, are done to obtain the electrode locations (SOZ) for each patient. Video-EEG monitoring was also reported for drug-resistant epilepsy patients as a pre-surgical evaluation.

The subdural electrodes (4-mm diameter and 10-mm distance) (UNIQUE MEDICAL Co, Tokyo, Japan) were put in and covered almost the whole surface on the FCD and the adjacent cortex. In patients with the bottom of sulcus (BOS) type of dysplasia, the surgeon dissected the cortical sulcus and implanted small electrodes on the vertical sulcus. The Neuro Fax digital video EEG system (NIHON-KODEN, Tokyo, Japan) were used to obtained the interictal iEEG data where the sampling rate of each patients was 2 kHz. The number of electrodes were determined for each patient based on an epileptologist's review during data recording where a positive label was assigned to a channel judged to a seizure onset electrode (SOZ) and a negative label was given to the rest of the channels (non-SOZ). Table 4.1 shows the summary of the interictal iEEG dataset from the eight patients.

TABLE 4.1: The summary of eight patient datasets of interictal iEEG with focal cortical dysplasia (FCD). M and F represent respectively the male and female.

Patients ID	Age and sex	Lesion site	Location	Pathology	Sampling frequency	No. of electrodes	No. of SOZ electrodes	Follow up	Engel
Pt1	5/F	Lt dorsal superior temporal gyrus	Cortical surface	Type 2B	2 KHz	60	3	3 years	IA
Pt2	39/F	Lt dorsal superior frontal gyrus	Bottom of sulcus	Type 2B	2 KHz	50	10	3 years	IA
Pt3	5/M	Lt cingulate gyrus	Bottom of sulcus	Type 2B	2 KHz	42	6	3.5 years	IA
Pt4	6/M	Rt dorsal middle frontal gyrus	Cortical surface	Type 2B	2 KHz	36	3	3.5 years	IA
Pt5	20/M	Rt middle frontal gyrus	Cortical surface	Type 2A	2 KHz	60	6	4.5 years	IIIA
Pt6	15/M	Lt superior parietal lobule	Cortical surface	Type 2B	2 KHz	70	7	5 years	IA
Pt7	32/M	Lt superior parietal lobule	Bottom of sulcus	Type 2B	2 KHz	70	10	5 years	IA
Pt8	25/M	Lt angular gyrus	Bottom of sulcus	Type 2A	2 KHz	76	16	5 years	IA

4.1.2 Focal cortical dysplasia (FCD)

In this analysis, we have used the focal cortical dysplasia (FCD) type data. FCD refers to the localized area of abnormal cerebral cortex that are most often associate to patients with epilepsy in both children and adults. For diagnosis of FCD, a wide spectrum of histopathology are used [91]. In FCD, the sites of interictal epileptiform discharges (IED) including spikes, polyspikes, sharp waves etc. appearance and seizure initiation are centered on imaging sites. Moreover, different studies are reported that, there are a high seizure suppression rate in FCD type 2 by excising the abnormal imaging sites.

4.1.3 Multiband analysis

In practice, the EEG time series exhibits nonstationary behavior with a variety of neurological events which may hold artifact that can reduce the system performance in a single-band approach. Therefore, a filter bank was applied to decompose an EEG signal into a set of multiple subband signals. [73], [92]. For more accurate detection of brain activities connected to the particular mental tasks, several EEG based studies [93]–[95] proposed the filter-bank

method that splitted the wide frequency ranges in narrow subbands. More specific, Higashi et al. showed a filter-bank approach in their MI-BCI studies to improve the performance of the system. They divided the 4–40 Hz frequency ranges into 6 subbands each of which has a bandwidth of 6 Hz [96]. Ang et al. decomposed the similar frequency ranges (4–40 Hz) into small subbands each of which has a bandwidth of 4 Hz [93]. In signal processing study, the choice of subbands should be as narrow as possible to achieve more accurate detection similar to these EEG-BCI [93]–[96]. Therefore, the choice of dividing the wide ranging frequency bands into narrow subbands actually relates to the system performance, real-time applications and reduction of system complexity [40], [94], [97].

4.1.4 Entropy based feature extraction methods

To extract features characterizing the complexities of a time-series from interictal iEEG signals, different entropy-based methods are available. Several epilepsy related studies has been reported [39], [79] that the combination of various entropy based feature extraction methods can improve the classification performance. Therefore, the eight entropy based feature extraction methods with a multi-band approach was chosen to extract features in this study. The details of the eight entropy measures used in this study are summarized in the following sections.

4.1.4.1. Approximate entropy

Pincus et al. first introduced the approximate entropy (APE) for measurement of regularity in the time-series [98]. It is broadly used in many fields of biomedical signal processing, such as EEG [99] and ECG signal analysis [100]. To calculate the APE from each segment, let us define a time series as

$x(i)$ of the n-th subbands S_n of each channel. The time series $x(i)$ can be defined $N-d+1$ vectors as $X(1), X(2), \ldots, X(N-d+1)$, where N is the signal length (in our case $N = 40000$ for each channel of a segment). Each vector $X(i)$, can be represented as:

$$X(i) = [x(i), x(i+1), \ldots, x(i+d-1)] \in \mathbb{R}^d, \tag{4.1}$$

where d is the embedding dimension and for each i, $1 \leq i \leq N-d+1$. APE is defined as:

$$APE(d, r, N) = \frac{1}{N-d+1} \sum_{i-1}^{N-d+1} \ln(C_i^d(r)) - \frac{1}{N-d} \sum_{i-1}^{N-d} \ln(C_i^{d+1}(r)), \tag{4.2}$$

where $C_i^d(r)$ is a correlation integral representing the probability of the vector $X(i)$, which remains similar to $X(j)$ within tolerance limit r. The $C_i^d(r)$ is defined as [101], [102]:

$$C_i^d(r) = \frac{1}{(N-d+1)} \sum_{j=1}^{N-d+1} \mathbb{I}\left(dist(X(i) - X(j)) \leq r\right), \tag{4.3}$$

where $\mathbb{I}(\cdot)$ is the indicator function and the $dist(\cdot)$ indicates the distance between two vectors $X(i)$ and $X(j)$. In our study, the value of the r parameter is set as the 0.2 times the standard deviation of the signal, and $d = 2$ [98].

4.1.4.2. Sample entropy

Sample entropy (Sp) is a modified version of approximate entropy (APE) and it is introduced to resolve the problem of APE [103]. The main disadvantage of APE is a biased estimate because of self-matches of templates. Sample entropy diminish the bias caused by the usege of the self matches in

the count of APE [103]. For a given time series $x(i)$, Sp is defined as:

$$Sp(d, r, N) = -\ln\left(\frac{A^d(r)}{B^d(r)}\right), \tag{4.4}$$

where

$$B^d(r) = \frac{1}{(N-d)}\sum_{i=1}^{N-d} C_i^d(r), \tag{4.5}$$

$$A^d(r) = \frac{1}{(N-d)}\sum_{i=1}^{N-d} C_i^{d+1}(r). \tag{4.6}$$

The $C_i^d(r)$ is defined as [101], [102]:

$$C_i^d(r) = \frac{1}{(N-d)}\sum_{j=1}^{N-d} \mathbb{I}\left(dist(X(i) - X(j)) \leq r\right), \tag{4.7}$$

where $X(i)$ is a vector induced from eq. (4.1) and $\mathbb{I}(\cdot)$ is the indication function to count the true condition number excluding the self-matches $\mathbb{I}\left(dist(X(i) - X(i)) = 0 \leq r\right)$ [98]. In this study, the parameters r and d were set to the similar to the approximate entropy.

4.1.4.3. Permutation entropy

Permutation entropy (PE) is a simple and robust method for estimating the complexity of a time series used for automated seizure prediction [104]. For a given time series $x(i)$, each vector $X(i) = [x(i), x(i+\tau), \ldots, x(i+(d-1)\tau)]$, where the d and τ are the embedding dimension and time lag, respectively. Let us define a permutation of $[1, 2, \ldots, d]$ by $\Pi = [j_1, j_2, \ldots, j_d]$ in such a way that $x(i+(j_1-1)\tau) \leq x(i+(j_2-1)\tau) \leq \ldots \leq x(i+(j_d-1)\tau)$. Then, we can define $\tilde{X}(i) = [x(i+(j_1-1)\tau), x(i+(j_2-1)\tau), \ldots, x(i+(j_d-1)\tau)]$. For the set of vectors $\{X(i)\}_{i=0}^{N-(d-1)\tau}$, the probability of each possible permutation Π_k $(k = 1, 2, \ldots, d!)$ can be introduced as $p(\Pi_k) = C(\Pi_k)/(N-(d-1)\tau)$, where N is the length of time series $x(i)$ and $C(\Pi_k)$ is the number of occurrences of

the order pattern Π_k. The PE can be defined as:

$$PE = -\sum_{k=1}^{d!} p(\Pi_k) \log_2 p(\Pi_k), \tag{4.8}$$

In this study, the parameters d and τ were set to 3 and 1, respectively.

4.1.4.4. Spectral entropy

Spectral entropies quantify the complexity of a time series based on the power spectrum [105]. Several studies have proposed the use of spectral entropy, including Shannon (Sh) and Reny's entropy (Ren), to characterize the seizure activities [105]–[107]. To obtain the power level for each frequency, the Fourier transform of the time series $x(i)$ is used. The normalization of the power p_f was estimated as:

$$p_f = \frac{P_f}{\sum P_f}, \tag{4.9}$$

where P_f is the power level of the frequency component. The entropies defined as Sh and Ren are estimated in follows [105]:

$$Sh = -\sum_f p_f \ln(p_f), \tag{4.10}$$

$$Ren\,(\alpha) = \frac{1}{1-\alpha} \sum_f \ln p_f^2, \tag{4.11}$$

where α is the order of Reny's entropy ($\alpha = 2$).

4.1.4.5. Phase entropy

Phase entropies are defined through a bispectrum known as higher order spectra [108]. The bispectrum of a time series $x(i)$ can be defined as:

$$B(f_1, f_2) = E[F\,(f_1)\,F\,(f_2)\,F^*\,(f_1 + f_2)], \tag{4.12}$$

where E represents the expectation operator of a random variable. The F is the Fourier transform of the time series $x(i)$ and F^* is its conjugate. The two types of phase entropy, S_1 and S_2, can be defined as:

$$S_1 = -\sum_k p_k \ln(p_k) \tag{4.13}$$

$$S_2 = -\sum_r q_r \ln(q_r) \tag{4.14}$$

where $p_k = \frac{|B(f_1,f_2)|}{\sum_{\Omega \in f_1,f_2} |B(f_1,f_2)|}$ and $q_r = \frac{|B(f_1,f_2)|^2}{\sum_{\Omega \in f_1,f_2} |B(f_1,f_2)|^2}$.

4.1.4.6. Tsallis entropy

Tsallis entropy is the generalized version of Shannon entropy and controls the trade off between the contributions from the tails and the main mass of the distribution [109]. Tsallis entropy is defined as [109]:

$$Ts = \frac{1 - \sum_f p_f^q}{q - 1}, \tag{4.15}$$

where p_f is the normalization of power computed from the eq. (4.9) and q is a real number, frequently called the entropic-index, that characterizes the degree of non-extensivity of the framework [109], [110]. In this study, we set $q = 2$.

4.2 Features selection

In machine learning approach, one of the challenging issues is the selection of prominent features from all the available feature space [95], [111]–[113]. The selection of entropy features could provide a more accurate classification with respect to the entire set of features. Sparse LDA (sLDA) is a recently advanced feature selection technique [114], [115], that disclose discriminant directions of a few variables instead of all the variables used in the standard

LDA [116], [117]. In our study, we have used sLDA to select the prominent entropy features. After extracting entropy features from an interictal iEEG segment, the entropies of n-th subband for a channel can be expressed as:

$$\mathbf{v}_n = [u_n^{(1)}, u_n^{(2)}, \ldots, u_n^{(D)}] \in \mathbb{R}^D \tag{4.16}$$

where $\mathbf{v}_n$ denotes the combination of entropies ($D = 8$) extracted from the n-th subband of a channel using the above feature-extraction methods. We can calculate the entropies for all channels with each segment and finally stacked all of the segments to form the training features $M_n \in \mathbb{R}^{H \times D}$, where $H = ch \times s$ such that ch and s are the total number of channels and segments, respectively. The sparse LDA criterion from the set of the training features M_n and class C_n for n-th subband is defined sequentially as [118]:

$$\left\{\hat{\theta}_n, \hat{\beta}_n\right\} = \arg\min_{\theta_n, \beta_n} \|C_n\theta_n - M_n\beta_n\|_2^2 + \delta \|\beta_n\|_2^2 + \delta_1 \|\beta_n\|_1$$
$$\text{subject to } \frac{1}{H}\theta_n C_n^T C_n \theta_n = 1 \tag{4.17}$$

where $\theta_n = (1, 1)^T$ is the initialization vector. The δ and δ_1 are tuning parameters used to achieve non-zero elements in each discriminative direction. By solving eq. (4.17), we will achieve the $\hat{\beta}_n = [\hat{\beta}_n^{(1)}, \hat{\beta}_n^{(2)}, \ldots, \hat{\beta}_n^{(D)}]^T$. The parameters δ and δ_1 were tuned such that $\hat{\beta}_n$ has G non-zero elements. Let us define the index of the $\hat{\beta}_n$ as $I_n = \left\{i, 1 \leq i \leq D | \hat{\beta}_n^{(i)} \neq 0\right\}$. The features $\tilde{\mathbf{v}}_n$ for n-th subband with a channel can be defined as:

$$\tilde{\mathbf{v}}_n = [u_n^{I_n(1)}, u_n^{I_n(2)}, \ldots, u_n^{I_n(G)}] \in \mathbb{R}^G \tag{4.18}$$

Finally, feature $\mathbf{V}^*$ is defined by concatenating features $\tilde{\mathbf{v}}_n$ of L subbands for a channel as:

$$\mathbf{V}^* = [\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \ldots, \tilde{\mathbf{v}}_L] \in \mathbb{R}^{LG} \tag{4.19}$$

After applying sLDA, the selected training features for L subbands can be written as:

$$\nu = \left\{ \mathbf{v}_F^{*(i_{in})}, \mathbf{v}_{NF}^{*(j_{in})} : i_{in} = 1, \ldots, I_{in}; j_{in} = 1, \ldots, J_{in} \right\} \tag{4.20}$$

The feature vector of the i_{in}-th sample of SOZ channel is denoted by $\mathbf{v}_F^{*(i_{in})}$ and the feature vector of the j_{in}-th sample of non-SOZ channel is denoted by $\mathbf{v}_{NF}^{*(j_{in})}$. Note that the dataset is generally imbalanced say $I_{in} << J_{in}$.

4.3 Imbalanced learning problem

In case of seizure onset zone detection, the number of non-SOZ channels illustrating the majority class is much higher than the SOZ channels (minority class). This imbalance class distribution can create several difficulties in standard machine learning approachs [119]–[121]. Therefore, imbalanced learning applications requires the modification of an imbalanced data set with some mechanisms in sequence to provide a balanced distribution. Several recent studies [119]–[122] have shown that a balanced data set with various base classifiers provides for improved classification performance in contrast to an imbalanced data set. In this section, we generate surrogate data using the adaptive synthetic (ADASYN) approach, which is one of the solutions used to solve the imbalanced learning problem. The balance set $\tilde{\nu}$ can be defined from the training feature set ν induced from eq. (20) as:

$$\tilde{\nu} = \left\{ \mathbf{v}_F^{*(i_{in})}, \mathbf{v}_{NF}^{*(j_{in})}, \mathbf{v}_F^{*(\tilde{i}_{in})} : i_{in} = 1, \ldots, I_{in}; j_{in} = 1, \ldots, J_{in}; \tilde{i}_{in} = 1, \ldots, \tilde{I}_{in} \right\}, \tag{4.21}$$

where, $I_{in} + \tilde{I}_{in} = J_{in}$. The following algorithm, proposed by He et al. [119], [120], is employed here to generate surrogate samples $\mathbf{v}_F^{*(\tilde{i}_{in})}$.

Step 1 Calculate the number of synthetic data examples that need to be generated for the entire focal class by:

$$\tilde{I}_{in} = (J_{in} - I_{in}) \times \beta \tag{4.22}$$

The β represents an arbitrary number in the range of 0 to 1 to specify the desired balance level after the synthetic data generation process. We set the β in eq. (4.22) to 1, which corresponds to fully balanced data [119].

Step 2 For each example $\mathbf{v}_F^{*(i_{in})}$ in the focal class, find the K-nearest neighbors according to the Euclidean distance and calculate the ratio $\Gamma_{i_{in}}$ as follows:

$$\Gamma_{i_{in}} = \frac{\Theta_{j_{in}}/K}{Z}, \quad i_{in} = 1, 2, \ldots, I_{in}, \tag{4.23}$$

where $\Theta_{j_{in}}$ is the number of samples in the K-nearest neighbor of $\mathbf{v}_N^{*(j_{in})}F$ that belong to the non-focal class and Z is a normalization factor such that $\Gamma_{i_{in}}$ is a distribution function ($\sum_{i_{in}} \Gamma_{i_{in}} = 1$).

Step 3 Determine the number of synthetic samples to be generated for each $\mathbf{v}_F^{*(i_{in})}$ in the focal class as:

$$g_{i_{in}} = \Gamma_{i_{in}} \times \tilde{I}_{in} \tag{4.24}$$

Step 4 Generate $g_{i_{in}}$ synthetic data samples for each sample of focal class using SMOTE algorithm [123] as:

$$\mathbf{v}_F^{*(\tilde{i}_{in})} = \mathbf{v}_F^{*(i_{in})} + (\tilde{\mathbf{v}}_F^{*(i_{in})} - \mathbf{v}_F^{*(i_{in})}) \times \delta \tag{4.25}$$

where $\tilde{\mathbf{v}}_F^{*(i_{in})}$ is a randomly chosen focal data example from the K-nearest neighbors ($K = 5$) for $\mathbf{v}_F^{*(i_{in})}$ and δ denotes the random number belonging to

TABLE 4.2: Confusion matrix for a two-class problem.

	Predicted positive	**Predicted negative**
Actual positive	TP: True Positive	FN: False Negative
Actual negative	FP: False Positive	TN: True Negative

[0,1]. The other parameters for ADASYN were used as default setting [119].

4.4 Methodological framework evaluation

4.4.1 Cross-validation design

To evaluate the proposed design, we have divided the data into training and test sets, which is the critical step due to an imbalanced number of SOZ and non-SOZ channels. To optimally divide the data into training and test sets, this study proposes k-fold cross-validation technique ($k = 10$) by dividing 90 segments into k subsets of equal size. Among the k subsets, one subset is used for testing and the remaining $(k - 1)$ subsets are used for training. As mentioned that, ADASYN was applied to balance the data in the training strage. The cross-validation process is then repeated k times and the result of a system is taken by averaging all the runs.

4.4.2 Performance measurement for segments

To measure the pergformance of our proposed design, a set of assessment metrics related to receiver operating characteristics (ROC) [124] graphs were used. Under the imbalanced learning case, the classification accuracy is not adequate as a standard performance measurement. [124]–[127]. Therefore, the representation of classification performance can be derived from the confusion matrix, as illustrated in Table 4.2. Based on this table, the evaluation metrics can be expressed as:

- Sensitivity (SEN) or recall:

$$SEN = \frac{TP}{TP + FN} \times 100\%, \tag{4.26}$$

where TP is the number of correctly detected segments from the total number of SOZ segments in the SOZ channels and FN indicates the number of correctly detected segments from the total number of non-SOZ segments in the non-SOZ channels.

- Specificity (SPE):

$$SPE = \frac{TN}{TN + FP} \times 100\%, \tag{4.27}$$

where TN is the number of correctly detected segments from the total number of non-SOZ segments in the non-SOZ channels and FP represents the number of incorrectly detected segments from the total number of non-SOZ segments in the non-SOZ channels.

- Precision or positive predictive value (PPV):

$$Precision = \frac{TP}{TP + FP} \times 100\% \tag{4.28}$$

- Fall-out or false positive rate (FPR):

$$FPR = \frac{FP}{TN + FP} \times 100\% \tag{4.29}$$

- $F_1\ score$ is the harmonic mean of preision and sensitivity defined as:

$$F_1\ score = 2 \cdot \frac{Recall \times precision}{Recall + precision} \tag{4.30}$$

4.4.3 Performance measurement for channels

. For each-fold cross validation, the sensitivity (SEN) and false positive rate (FPR) of the channels were calculated do with each threshold values (in our case, zero to maximum number of identified SOZ segments for each fold). After acquiring SEN and FPR of all channels we calculated the AUC by using trapezoid rule [128] and average of all the folds to attain the final results.

4.5 Experimental result

To measure the performance of the system, firstly the AUC of three different methods such as filter-bank approach (FbA), FbA with ADASYN (FbA/ADA), and FbA with feature selection and ADASYN (FbA/FS/ADA) were compared with 10-fold cross validation. Secondly, based on the performances of AUC, the optimum one was selected and used to simulate the result of eight epilepsy patients. The three different proposed algorithms were shortly brief as follows:

Algorithm 1- (FbA): In this algorithm, L (L=10) bandpass filters were implemented using a third-order Butterworth bandpass filter to subband the high-frequency components (100–600 Hz) in interictal iEEG signal. Then eight entropy based feature extraction methods were applied to extract features from each subbands. After that, the extracted features were input to SVM classifier with a 10-fold cross-validation for classifying the SOZ and non-SOZ segments.

Algorithm 2-(FbA/ADA): In this case, subbanding and feature extraction procedure were performed in the same way as in **Algorithm 1**. ADASYN method was applied in the training stage of the SVM classifier for each cross-validation to deal with the the imbalanced learning problem.

Algorithm 3-(FbA/FS/ADA): The N bandpass filters and feature extraction were accomplished in the same way as in **Algorithm 1**. In this algorithm, sparse LDA was used to select the prominent entropy features among all eight entropy features. Finally, the selected entropy features were used for classifying SOZ segments with SVM classifier. As mention that, the ADASYN method was also used the same purpose described in **Algorithm 2**.

4.5.1 Effect of feature selection.

In this studies, we used eight entropy based feature extraction methods to extract features from each subband defined in eq. (4.16). Among eight entropy features, we select the prominent entropy features by using sLDA weights from the training set showed in eq. (4.17) based on non-zero weights from each subband. In our study, we set the sparsity parameters $\delta = 3$ and δ_1 (Pt1: $\delta_1 = -5$; Pt2: $\delta_1 = -5$; Pt3: $\delta_1 = -3$; Pt4: $\delta_1 = -5$; Pt5: $\delta_1 = -3$; Pt6: $\delta_1 = -4$; Pt7: $\delta_1 = -6$; Pt8: $\delta_1 = -3$) based on the training set to improve the results, where the absolute value of δ_1 corresponds to the desired number of variables. Fig. 4.2 shows the colormap of sLDA weights of each entropy features for all eight patients where vertical-axis represents the sLDA weight for each subband and horizontal-axis represent the weights of each entopy features. The figure indicates that the features with non-zero weights are more significant. To justify this hypothesis, AUC area under the ROC curve was derived with 10-fold cross-validation from individual entropy features, as well as the average weights across N subbands, were estimated, as illustrated in Fig. 4.3. This figure shows the relationship between the weights of entropies and the AUC of individual entropies suggesting that the entropies with non-zero weights may improve the design performance. For performance evaluation, entropy features corresponding to non-zero weights were

selected for the purpose of SOZ and non-SOZ classification.

4.5.2 Performance analysis with different cases.

To evaluate the performance in three algorithms
(FbA, FbA/ADA, and FbA/FS/ADA), the AUC was performed for eight patients shown in Fig. 4.4. In case of imbalanced learning , the AUC is equivalent to the possibility of ranking a randomly selected positive instance higher than a randomly chosen negative instance [124].

In this study, we set all parameter for ADASYN following the study [119] to balance the training features. The AUC for the algorithm FbA/ADA with feature selection exhibits superior results for all eight patients. The reason for lower performance using the FbA method is that the high degree of imbalance distribution between the minority (focal segments) and majority (non-focal segments) class may provide biased decision boundary used in SVM training. In the test of the statistical significance of the methods, the result of Friedman's ANOVA showed a significant main effect on AUC ($p < 0.05$). Performing a tukey-kramer-based post-hoc test, the method using FbA/FS/ADA achieved significantly higher AUC across all eight patients than the other methods (FbA vs FbA/ADA: $p < 0.001$; FbA/ADA vs FbA/FS/ADA: $p < 0.001$).

4.5.3 Results with SOZ segments-spotting.

Automatic identification of individual SOZ segments of eight epilepsy patients based on the optimal algorithm (FbA/FS/ADA) is provided in this section. The result showed that the selection of useful entropy features with combination of oversampling methods can significantly improve the performance of proposed SOZ detection design. However, we are the first time to used high frequency components (100–600 Hz) to detect SOZ segments from

TABLE 4.3: Experimental results for segment detection using our proposed optimal method (FbA/FS/ADA).

Patient ID	SEN [%]	SPE [%]	Precision [%]	Fall-out [%]	F-score
Pt1	52.96	97.13	49.31	2.87	0.51
Pt2	27.33	92.83	47.88	7.17	0.34
Pt3	51.67	76.98	27.22	23.03	0.36
Pt4	62.96	86.19	29.31	13.80	0.40
Pt5	61.11	95.62	60.77	4.39	0.61
Pt6	96.82	91.83	56.85	8.16	0.72
Pt7	65.33	94.93	68.21	5.07	0.67
Pt8	86.31	99.33	97.18	0.67	0.91
Mean	**63.03**	**91.85**	**54.59**	**8.15**	**0.57**

interictal iEEG signal. Like other different HFOs- and low frequency-based related studies [4], [30], [51], [129]–[133] the performances were measured in terms of sensitivity, specificity, precision, fall-out, F-score shown in Table 4.3. From this table it has been observed that, the proposed method achieved the highest performance for localizing individual segments for the adult patients Pt5 (SEN: 61.11%; Fall-out: 4.39%), Pt6 (SEN: 96.82%; Fall-out: 8.16%), and Pt8 (SEN: 86.31%; Fall-out: 0.67%). According to the different studies [134]–[136], we also consider the positive likelihood ratios (PLRs) to evaluate the performance of our proposed design framework. The acquired PLRs for the adult patients (Pt5, Pt6, and Pt8) are 13.92, 11.16, and 78.46, respectively. For the deep sheeted patients (Pt2 and Pt7), the surgeon implanted the small electrodes vertically on the sulcus. The sensitivity and Fall-out of Pt2 (SEN: 27.33%; Fall-out: 7.17%) and Pt7 (SEN: 65.33%; fall-out: 5.07%) and the PLRs are 3.81 and 12.88 for patients Pt2 and Pt7, respectively. In case of pediatric patients, the sensitivity and fall-out are Pt1 (SEN: 52.96%; Fall-out: 2.87%), Pt3 (SEN: 51.67%; Fall-out: 23.03%), and Pt4 (SEN: 62.96%; Fall-out: 13.80%). For Pt1, Pt3, and Pt4, the positive likelihood ratios (PLRs) are 18.45, 2.24, and 4.56 respectively.

4.5.4 Results with channel identification.

Fig. 4.5 shows the graphical representation with the detection of the SOZ and non-SOZ segments, which assist the specialist of epilepsy in two ways: (1) to visualize the detection of the SOZ and non-SOZ segments over duration of the interictal iEEG, and (2) the number of detected SOZ segments corresponding to the SOZ and non-SOZ electrodes. This provides effective information about the active electrodes, that are closely located to the SOZ.

From this figure, the y axis in the color map (left) illustrates the electrodes and the x-axis indicates the segment index. Each yellow spot in the color map indicates the detected SOZ segments. For each patient, the right side of the color map represent the number of detected SOZ segments (horizontal-axis) in each channel (vertical-axis) in which a group of bars (yellow) indicate the SOZ channels and black bars without color represent the non-SOZ channels. It is observed from the Fig. 4.5 that, a sharp yellow spotted areas are clearly visible for each SOZ electrode for the patients Pt1, Pt5, Pt6, Pt7 and Pt8. The detected SOZ segments (yellow spot) for the patients of Pt2, Pt3 and Pt4 are distributed through the non-SOZ electrodes. Fig. 4.6) shows the receiver operating characteristic (ROC) curve with the AUC value for each patient that are measured across all possible thresholds based on localized SOZ segments in the SOZ and non-SOZ channels (see in Fig. 4.5)

4.5.5 Computational time

The average computational time for each entropy with 10 subbands are measured using Python on iMac Pro (with Intel Xeon W processor and 128 GB RAM). Note that the average results are estimated with 100 runs at the testing phase to detect a single segment. Table. 4.4 shows the mean computational time (in seconds) at each entropy with 10 subbands. It is observed that, the phase entropy requires the highest computational time. The entropy with

TABLE 4.4: The mean computational time (s) for each entropy with 10 subbands.

Methods	**APE**	**PE**	**Sh**	**Sp**	**Ts**	**Phase (S1 and S2)**	**Ren**	**Total Time (s)**
time (s)	12.40	0.032	0.010	12.40	0.008	31.66	0.008	**56.51**

Ts, Ren, Sh, and PE requires shorter time compared to others. For a single segment test, the average computational time with eight entropies and 10 subbands are 56.51 s.

4.6 Discussion

According to the clinical guidelines involved to epilepsy surgery, the epilepsy surgeon need to consider implanting the intracranial electrodes to observe the seizure onset zone (SOZ), irritative zone, and symptomatic zone before the epileptic focus resection. The removed area of the brain through the epilepsy surgery includes epileptic zone (SOZ, a part of irritable zone and symptomatic zone). To determine the epileptic SOZ electrodes, the epilepsy specialist need to analyze and label 3 to 7 days iEEG data that depend on the patients conditions. In our proposed patient-dependent design, to detect the SOZ electrodes, we have used only label 30-minutes labeled interictal iEEG data to localize the epileptic SOZ electrodes. Such type of detection or estimation of SOZ from short period of interictal recording provide epileptologists a great assistance and can increase the number of iEEG analysis for patients with intractable epilepsy.

There are several epilepsy related studies that has been reported in [30], [132]. Most of these studies have used Bern-Barcelona and Bonn EEG datasets for classifying epilepsy seizure or epileptic focus. For example, Mursalin et al. presented an automated epileptic seizure detection approach with improved correlation-based feature selection and random forest classifier (RFC) [137]. They used Bonn datasets and the average accuracy of their study was

of 98.44%. About the the Bonn datasets, it consists of five EEG datasets represented as Set A (normal: healthy awake and eyes open), Set B (normal: healthy awake and eyes closed), Set C (epileptic: interictal), Set D (epileptic: interictal), and Set E (epileptic: ictal) with 100 single-channels and the time duration of each channel was 23.6 s.

In another study [138], wavelet packet entropy and hierarchical EEG classification were introduced for classifying normal vs ictal EEG and the average accuracy was 99.44%. A method based on discrete wavelet transforms (DWT) using entropy features was proposed, leading to a classification accuracy of 84% using k-nearest neighbor (kNN), probabilistic neural network (PNN), fuzzy classifier, and least squares support vector machine (LS-SVM) [139]. There are also other EEG and iEEG datasets such as Friburg [133], CHB-MIT [4], Children's Hospital Boston datasets [140] etc. were used to detect the epileptic events based on different machine learning approaches.

Recently, Ullah et al. have used the Bonn datasets to detect the epilepsy seizure based on deep learning approach and the classification accuracy of their proposed study was 99.1% accuracy [38]. A similar dataset was used to design a deep convolutional neural network (CNN) with 13-layer for categorizing the normal, preictal, and seizure class and acquired an average accuracy of 88.7%, a specificity of 90% and a sensitivity of 95% [141]. Basically, the deep-learning based systems have improved the performance compared to simpler classier (k-NN, SVM etc.), but to showed the remarkable performance of these deep learning based system needs a large amount of training data. On contrary to the deep learning, the simple classifier method (SVM) is easy to understand and provides consistent performances. In another studies, Itakura et al. [40] used Bern-Barcelona dataset for identifying epileptic focus with average accuracy was 86.89%. However, the limitation all of the above studied was that, they used only lower frequency bands (0.5–150 Hz) for limited pairs of electrodes with well balanced problems.

For detecting epileptogenic zone or seizure onset zone, several studies [46], [142]–[144] have shown that, HFO may occur during ictal, preictal, and interictal states and the rate of HFOs tends to be higher in those zones [48], [144]. To identify the HFOs, several studies have proposed different methods such as artifact rejection, estimating the energy of the signal using root mean square (RMS) amplitude, short-time line-length or others [145]–[148]. For automatic detection of HFOs, Jrad et al. [149] have proposed multi-class SVM in depth-EEG signals. For performance evaluation, they used sensitivity and false discovery rate (FDR) and their acquired average result with five drug-resistant epilepsy for ripple (sensitivity: 81.1% and FDR: 30.2%) and fast ripple (sensitivity: 74.6% and FDR: 6.3%). Johansen et al. have proposed CNN methods for detecting spikes as well as HFOs and an average AUC with five epilepsy patients was 0.94. Zuo et al. [51] introduced the convolution neural network based method for detecting the HFOs in ripple and fast ripple separately and attained average results with sensitivity (77.04% and 83.23% for ripples) and specificity (72.27% and 79.36% for fast ripples) compared their study to four traditional automated methods proposed in the RIPPLELAB toolbox [52]. Recently another study [150] proposed HFOs identifying system with the combination of short-time energy (STE) and CNN classifier. For system performance measurement, they used sensitivity and FDR evaluation matrix and compared their proposed method with three related existing studies [52], [134], [149]. They used five adult patients and their acquired average accuracy for ripple (sensitivity: 81.1% and FDR: 30.2%) and fast ripple (sensitivity: 74.6% and FDR: 6.3%).

However, all of the above HFOs related epilepsy studies have mainly concentrated on the detection of HFOs in ripple and fast ripple iEEG data and the performance evaluation metrics of their proposed studies were usually used based on their balanced or imbalanced problems. In addition, for HFOs-related studies to identify the feasible seizure onset electrodes require

the long-time iEEG data to compute the baseline. Compared to the above HFOs-related studies, we used only 30-minutes of interictal iEEG data and combined the ripple and fast ripple bands together with the multi-band fashion to detect the channels related to SOZ . The average sensitivity, specificity, and fall-out of our proposed design for individual segment detection with eight patients was 52.70%, 90.75%, and 9.24%, respectively. The average AUC for identifying SOZ channels for all eight patients (Pt1: 0.90 , Pt2: 0.79 , Pt3: 0.71, Pt4: 0.79, Pt5: 0.96 , Pt6: 0.94 , Pt7: 0.81 , Pt8: 99) was 0.86.

In order to accomplish a more effective methodological framework for real-life applications, we have considered further improvements in the following directions. First, we used 30-mins of interictal iEEG signal for detecting SOZ electrodes using entropy based feature extraction methods with 10-fold cross validation. The computational time for some prominent entropy features are high. So, we can extend our work by considering the computational time as well as real world time series forecasting way with the improvement of design performance. Secondly, Islam et al. [95] reported that the selection of prominent operational subbands can significantly improve the framework performance. Therefore, the possible extension of this study is to detect the most significant subbands in the high-frequency components, which may further improve our design performance in the future. Finally, patient-independent design could be one of the best solutions for future study. Patient-independent design is more practical and real life implementation, because in this design no need any label data. However, the problem is very challenging due to very different locations of electrodes and subjects-specific nature of epilepsy events. Thus, there are several avenues for further research to model the design framework with feature-extraction and classification.

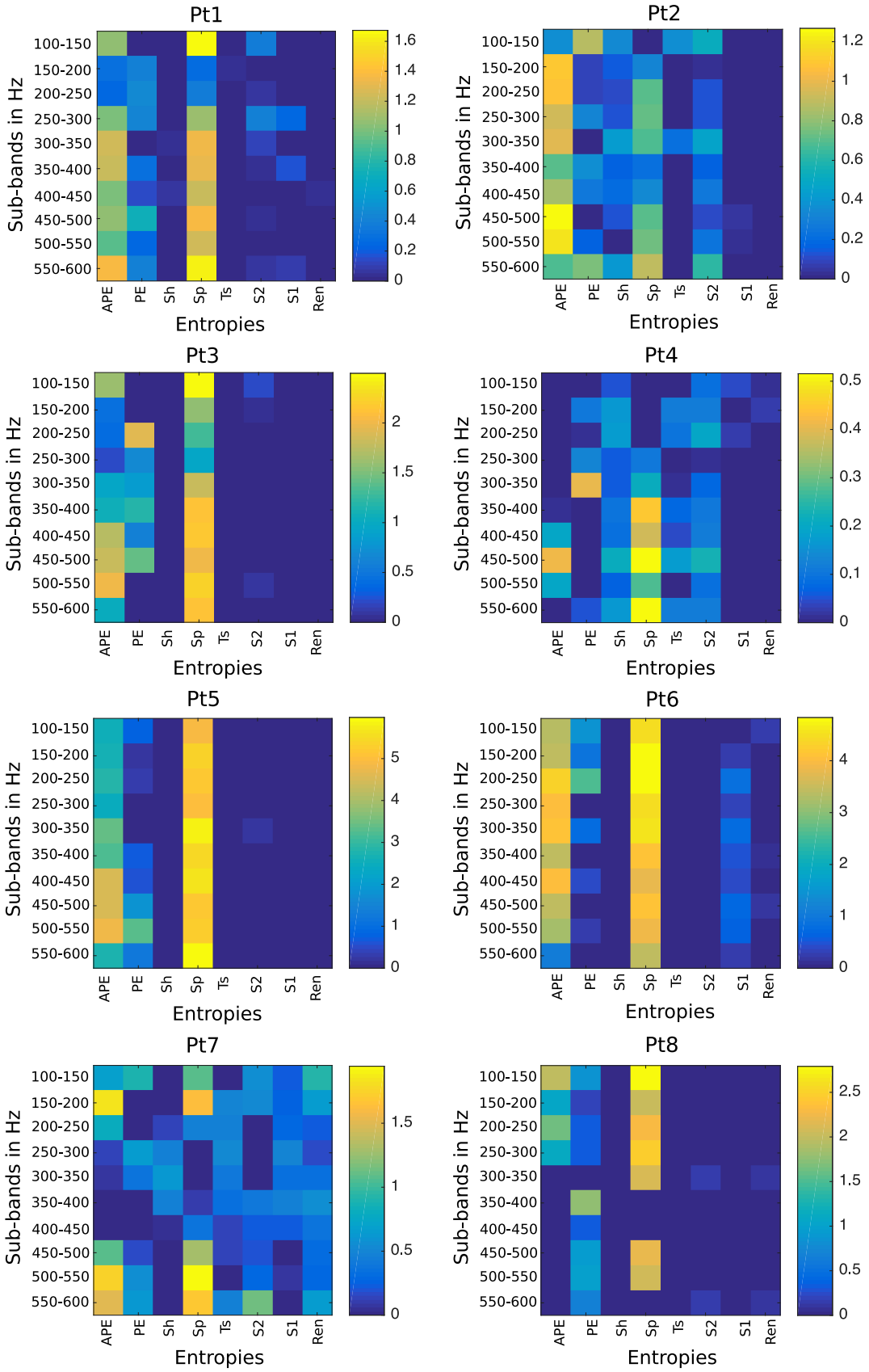

FIGURE 4.2: The color map representing the sLDA weights of the entropies with each subband for eight patients.

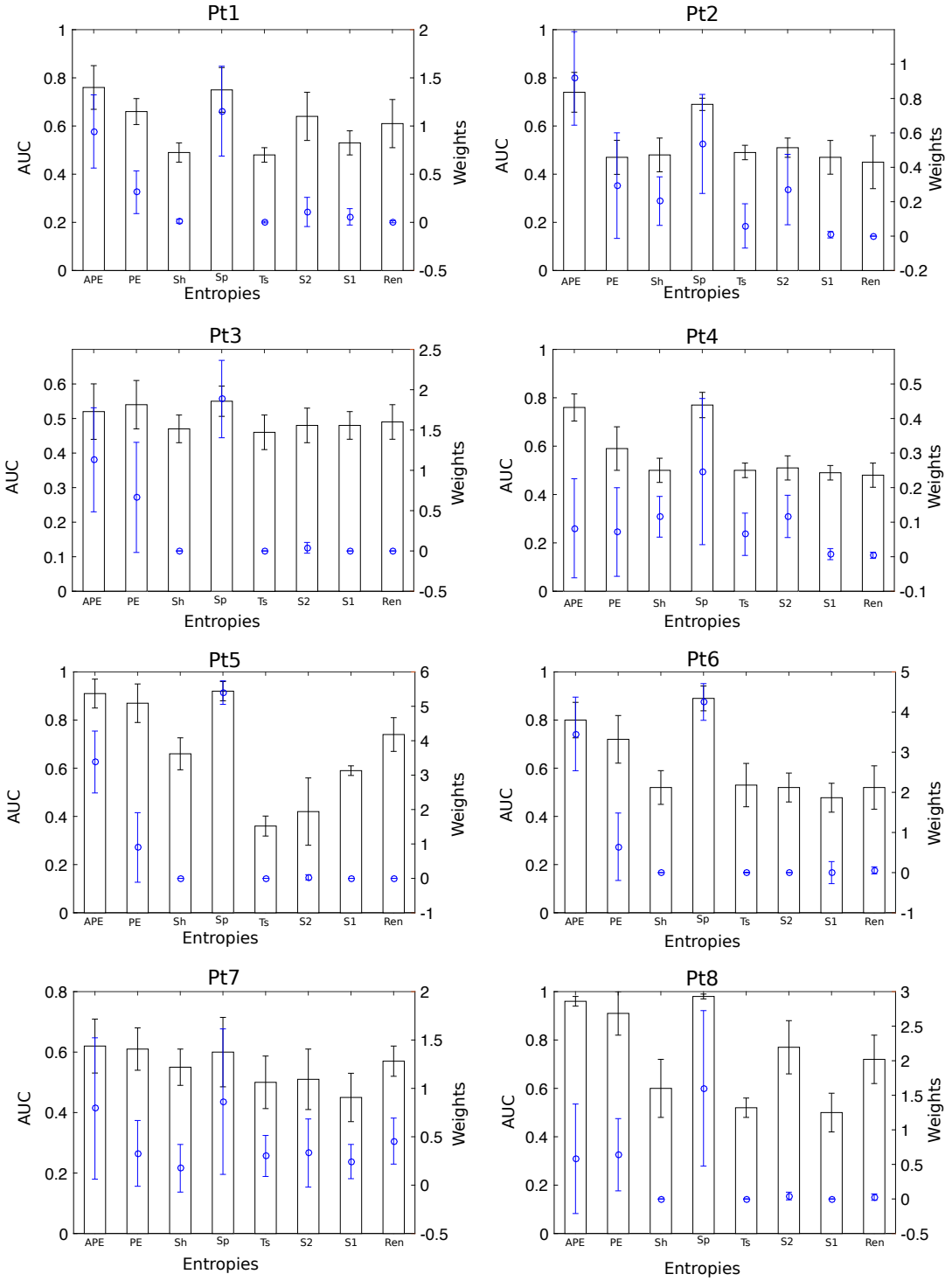

FIGURE 4.3: The bar diagram represent the average AUC for each entropy features and blue color circle inside bar indicates the average sLDA weights across subbands.

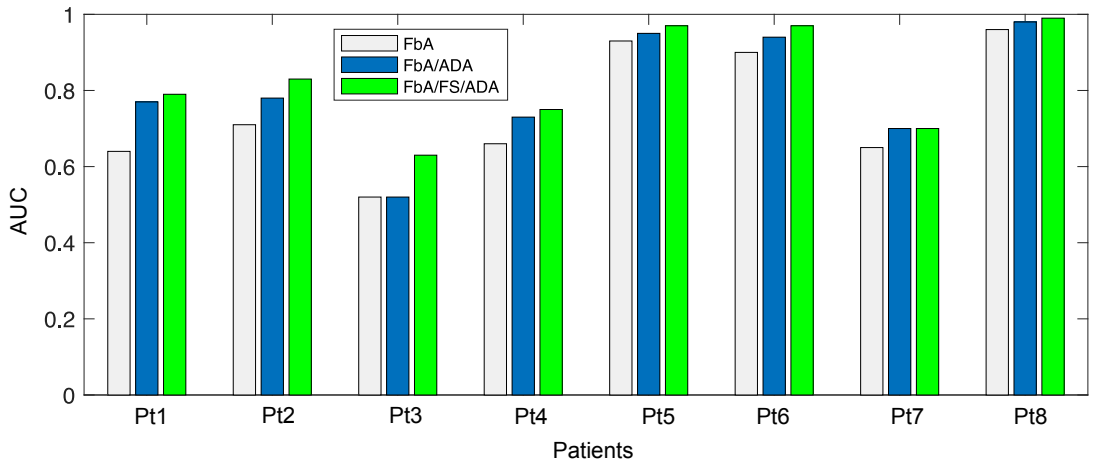

FIGURE 4.4: Average AUC with three different methods for all eight patients

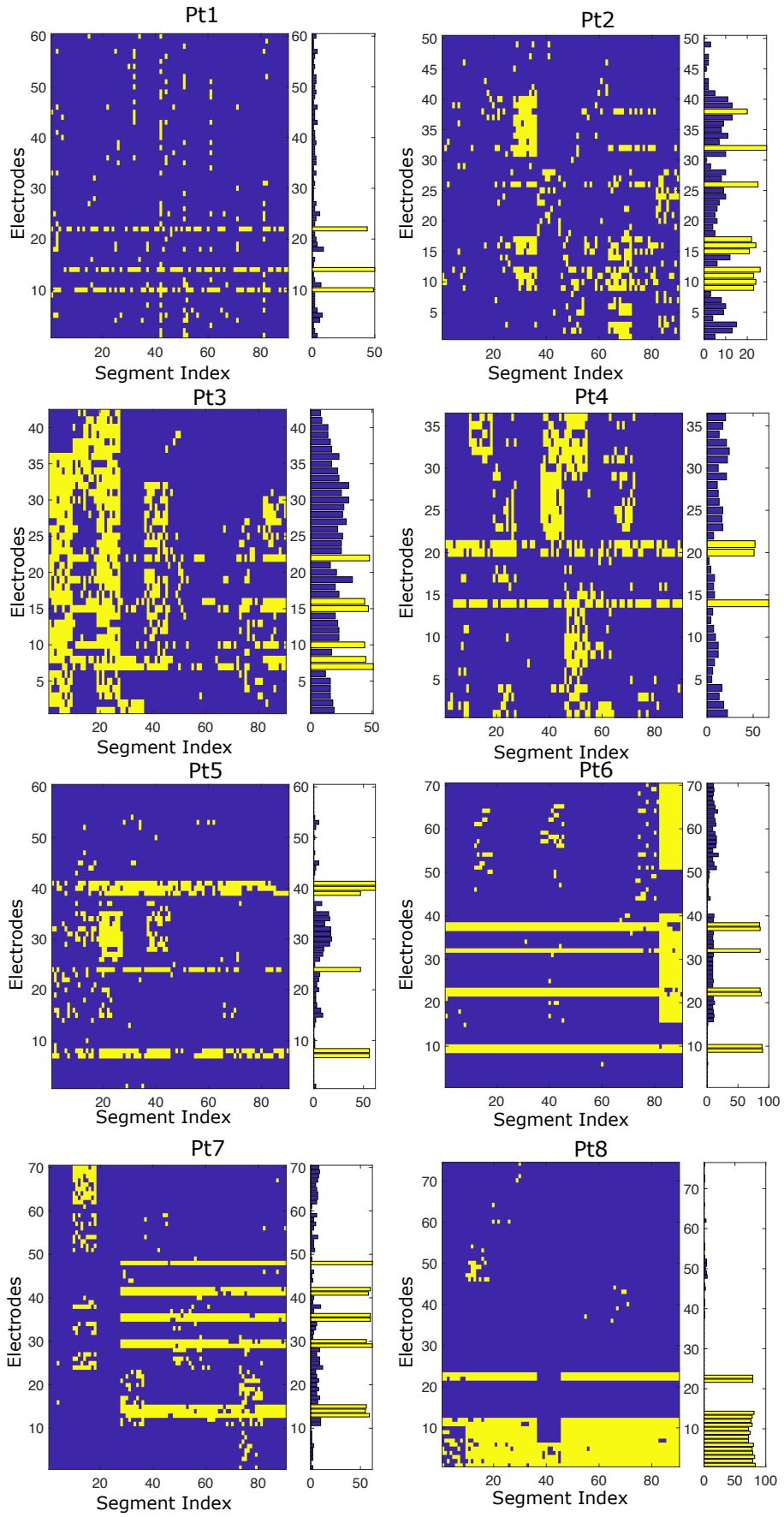

FIGURE 4.5: Color map representing the detection of SOZ segments (yellow spots) with respect to channels for the eight patients. The bar with each color map indicates the detected SOZ (yellow) and non-SOZ (black) electrodes with number of detected SOZ segments.

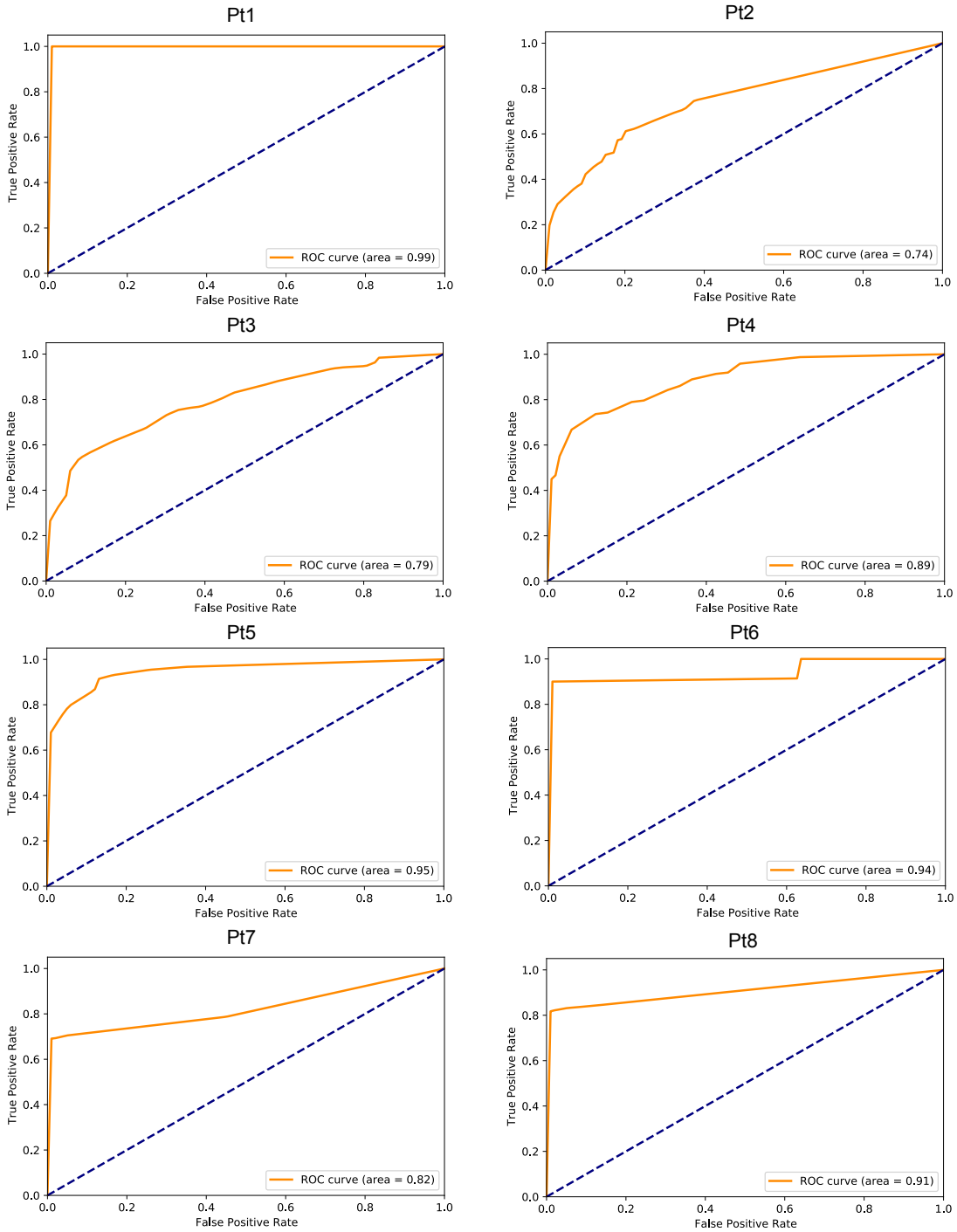

FIGURE 4.6: Average AUC with 10-fold cross-validation for identifying SOZ electrodes.

Chapter 5

Statistical Feature based SOZ Localization in Ripple and Fast Ripple Bands of Interictal iEEG

In Chapter 4, we have discussed about the high frequency component (HFC) based SOZ detection framework with entropy features. These entropy based study has suggested that, the entropy features can efficiently identify the SOZ electrodes in high frequency components. However, to design a computer aided solution, the entropy features has some shortcomings. Firstly, the performance of entropy estimation strongly depends on appropriate parameter selection [54]. For entropy measures, some of them use three or more than three parameters and the parameter choice of these entropy features depends on the length of test data. So, to design the method, there are many probable combination of parameter based on the data [54], [55]. Secondly, due to large sample point, some high performance entropy measure has higher computational cost. Thirdly, the detection of SOZ and non-SOZ segments was performed on hard thresholding based SVM classifier.

To address these problems, we hypothesized that simple statistical features are effective in identification of SOZ electrodes, since the statistical features used before in other context of epilepsy related study [31]–[33], [56],

[57]. Gotman [57] introduced the first extensively used method of seizure identification by simple calculations of amplitude from EEG signals. For automatic identification of epileptic seizure Khan et al. proposed a wavelet-based method from intracerebral electroencephalogram [33]. In their proposed system, they have used several feature extraction method such as energy, relative amplitude (RA), coefficient of variation (CV) to detect the seizure and non-seizure characteristics. Recently, Li et al. [31] has proposed coefficient of variation (CV) and fluctuation index (FI) based feature extraction method for automatic detection of epileptic seizure from ictal EEG signal. To enhance the performance of the system, Hassan et al. used ellipse area of second-order difference plot (SODP), CV, and FI as a feature-extraction method for seizure detection [32]. The above discussed studies collectively used the statistical features and they only focused to identify if an EEG segment was a seizure or non-seizure.

In this study our contributions are as follow: firstly, we have proposed twelve feature extraction methods including nine statistical features and three information theoretic entropy based methods to identify the SOZ electrodes. To the best of our knowledge, we have first time used the combination of these feature-extraction methods in HFCs (ripple and fast ripple bands) for detection of the SOZ electrodes. Secondly, mutual information (MI) based data-driven grid-search method has developed to select the prominent bands and features jointly. The selection of appropriate bands and features jointly in epilepsy related activities may improve the performance of the methods and still not reported in prior SOZ detection studies. To identify the possible SOZ electrodes, we have compared different methods with LightGBM and SVM classifies and select the optimal one. The identification was accomplished based on the scoring for every segment of channels calculated with the LightGBM algorithm.

5.1 Proposed design architecture

The block diagram of our proposed design shown in Fig. 5.1. According to this figure, the proposed design can be explain in the following way:

- **Data acquisition:** It includes description of datasets and its collection process. In this study, we have used eleven patients data collected from Juntendo University Hospital. The detailed description about this dataset in shown in Sec. 5.1.1.

- **Data preprocessing:** Data segmentation and filter bank analysis are included in this step. The discussion about data segmentation and filter bank analysis in Sec. 5.1.2.

- **Feature extraction:** In this step, we have used twelve statistical feature extraction methods to extract features from each subbands. The detail description about the feature extraction methods are introduced in Sec. 5.1.3.

- **Band and feature selection:** After extracted features from each subband, we have jointly select the prominent subbands and features using mutual information (MI) based grid search methods. The band and features selection part are broadly discussed in Sec. 5.1.5.

- **Data balancing:** Since our analyzed data are highly imbalanced that can create several difficulties in standard machine learning approach. To solve this problem, in this step we have used adaptive synthetic oversampling approach (ADAYSN) that were introduced in Chapter 4.

- **Classification:** Finally in this step, we have used two classifier such as SVM and LightGBM classifier to classify a channel whether a channel is SOZ or non-SOZ. This part is discussed in Sec. 5.1.6.

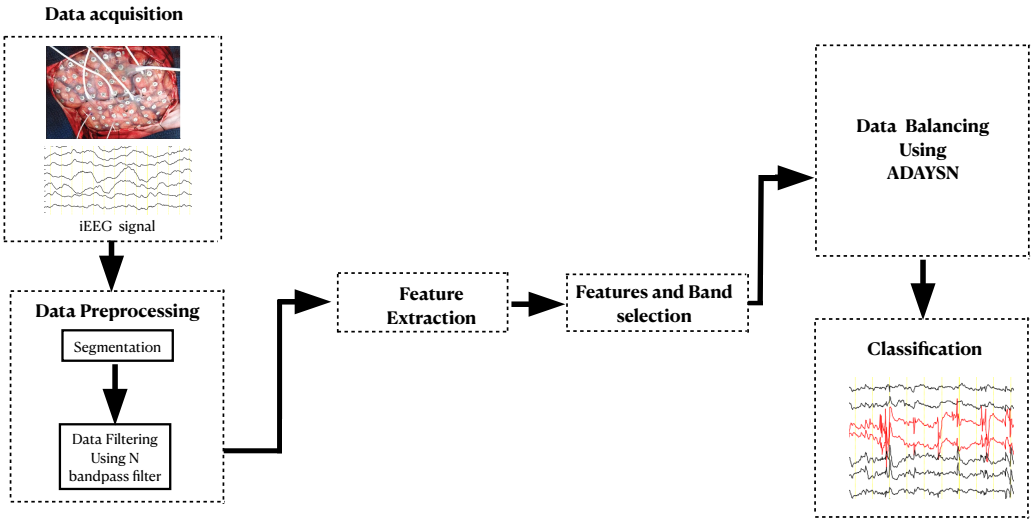

FIGURE 5.1: Block diagram of proposed design.

5.1.1 Dataset

In this study, we have used eleven patients data such as Pt1, Pt2, Pt3, Pt4, Pt5, Pt6, Pt7,Pt8, Pt9, Pt10 and Pt11 that are collected from Juntendo University Hospital, Tokyo, Japan. The dataset jointly approved by ethics committee of Juntendo University Hospital and Tokyo University of Agriculture and Technology. For data recording, the epilepsy surgeon implanted platinum subdural grids (UNIQUE MEDICAL Co., Tokyo, Japan) with 4-mm diameters and 10-mm distances for the cortical surface and platinum strip electrodes (UNIQUE MEDICAL Co., Tokyo, Japan) with 3-mm diameters and 5-mm distances for the vertical direction and the bottom of the cortex. All patients are long term interictal intracranial iEEG data with temporal lobe epilepsy caused by focal cortical dysplasia (FCD). The datasets includes six pediatric patients and five adult patients of ages between 5 to 39 years. Among eleven patients, the sampling frequencies of eight patients are 2 kHz and three patients are 1 kHz. Table 5.1 shows the summary of eleven patients data that were recorded by using the neuro fax digital video EEG system (NIHONKODEN, Tokyo, Japan) with length several days. The datasets have several

TABLE 5.1: The summary of interictal iEEG data for eleven patients with focal cortical dysplasia [151]. The male and female are indicated as M and F.

Patients ID	Age and sex	Lesion site	Location	Pathology	Sampling frequency	No. of electrodes	No. of SOZ electrodes	Follow up	Engel
Pt1	5/F	Lt dorsal superior temporal gyrus	Cortical surface	Type 2B	2 KHz	60	3	3 years	IA
Pt2	39/F	Lt dorsal superior frontal gyrus	Bottom of sulcus	Type 2B	2 KHz	50	10	3 years	IA
Pt3	5/M	Lt cingulate gyrus	Bottom of sulcus	Type 2B	2 KHz	42	6	3.5 years	IA
Pt4	6/M	Rt dorsal middle frontal gyrus	Cortical surface	Type 2B	2 KHz	36	3	3.5 years	IA
Pt5	20/M	Rt middle frontal gyrus	Cortical surface	Type 2A	2 KHz	60	6	4.5 years	IIIA
Pt6	15/M	Lt superior parietal lobule	Cortical surface	Type 2B	2 KHz	70	7	5 years	IA
Pt7	32/M	Lt superior parietal lobule	Bottom of sulcus	Type 2B	2 KHz	70	10	5 years	IA
Pt8	25/M	Lt angular gyrus	Bottom of sulcus	Type 2A	2 KHz	76	16	5 years	IA
Pt9	38/F	Rt supramarginal gyrus	Surface and vertical cortex	Type 2B	1 KHz	56	5	5.5 years	IIA
Pt10	14/F	Rt inferior frontal gyrus	Cortical surface	Type 2B	1 KHz	60	10	5.5 years	IC
Pt11	13/M	Lt angular gyrus	Surface and vertical cortex	Type 2B	1 KHz	68	2	5 years	IA

characteristics including patient ID, age and sex, lesion site, pathology, location, sampling frequency, number of electrodes, number of seizure onset zone (SOZ) electrodes, follow up, and Engel epilepsy surgery outcome scale. For analysis, epileptologists chosen the sleep stage iEEG without motion artifact. After selecting the data, they assigned label for each patients. The positive labels were given to the SOZ electrodes and negative labels were assigned to the rest of electrodes. After that, these label's data were used to design our machine learning based AI solution. From Table 5.1, IA Engel's class means the seizure free outcome. Mean follow up for all patients period was (4.9 ± 1.0) years. Seizure outcomes were evaluated using Engel's classification at the last visit to the outpatient center. All the patients are informed and they have signed the consent paper before recording of iEEG data.

5.1.2 Segmentation and filter bank analysis

To design the framework, we have used 1h (one hour) long interictal iEEG data, which were labeled by epieptilogist. As a poreprocessing step, we have divided this one hour multichannel iEEG data into 20-seconds segments and hypothesized that all of the segments of SOZ channels are SOZ segments whereas in the non-SOZ channels all of segments are non-SOZ. After splitting the data into segments, in each segment we have applied a third-order Butterworth bandpass filter to extract the high-frequency components from interictal iEEG.

Suppose that, S_n is the extracted subband components of iEEG signal using third-order Butterworth bandpass filter, where $n = 1, 2, \ldots, L$. In our case, the total number of subbands L was 10 for eight patients with sample frequency of 2 kHz and 7 for three patients with a sample frequency of 1 kHz. In other words, the cut-off frequencies were from 100–450 Hz for patients with 1 kHz sample frequency and the cut-off frequencies for patients with 2 kHz were from 100–600 Hz, which cover ripple and fast ripple bands in HFOs.

5.1.3 Feature extraction methods

To design a computer aided solution, feature-extraction is the important step. In this analysis, we have used twelve feature extraction methods including nine statistical features and three information-theoretic features to extract features from each subband. In different bio-signal processing reserach including iEEG, EMG, EEG etc., the following feature-extraction methods were proposed in different times. To extract the features from each epoch of intracranial EEG signals, let us define each channel of n-th subband as $\mathbf{x}$, which can be represented as $\mathbf{x} = [x_1, x_2, \ldots, x_N]$, where N is the length of $\mathbf{x}$. We

have used the combination of following features extraction methods that as introduced in the what follows.

5.1.3.1. Coefficient of variation

The coefficient of variation (CV) was used to analyze the characteristics of iEEG signal in different epilepsy studies [31]–[33], [56] that measure the dispersion of data. The CV indicates the ratio of standard deviation to the mean and it provides the information about the variation in any signal amplitude. For $\mathbf{x}$, the coefficient of variation can be defined as [31], [32]:

$$CV = \frac{\sigma}{\mu}, \tag{5.1}$$

where μ and σ represent the mean and standard deviation of the $\mathbf{x}$ computed as:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i, \tag{5.2}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2}, \tag{5.3}$$

5.1.3.2. Fluctuation index

In different epilepsy studies [31], [32], [152], the fluctuation index (FI) is used for measuring the intensity of signal amplitude changes. For $\mathbf{x}$, it can be defined as [31], [32], [56], [152]:

$$FI = \frac{1}{N-1}\sum_{i=1}^{N-1} \left| x_{i+1} - x_i \right|, \tag{5.4}$$

5.1.3.3. Variance

The variance (Var) of a signal refers to measuring how distant the amplitudes of the signal are expanding from their mean value. According to the different studies, it is also used to categories the seizure and non-seizure of a signal by using the Bonn dataset [56], [153]–[155]. The variance of the x can be expressed as:

$$Var = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2, \tag{5.5}$$

5.1.3.4. Root mean square

The root mean square (RMS) is one of the popular feature extraction method that is used to identify the HFOs in ripple and fast ripple bands of interictal intracranial EEG signals [52], [56], [156], [157]. In mathematics, RMS can be represented as:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}{x_i}^2}, \tag{5.6}$$

5.1.3.5. Difference absolute standard deviation

In different biomedical signal processing studies, the difference absolute standard deviation ($DASD$) is used as a popular statistical feature extraction methods [56], [158], [159]. $DASD$ the resemble to root mean square feature and can be represent as:

$$DASD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(x_{i+1} - x_i)^2}, \tag{5.7}$$

5.1.3.6. Mean absolute value

The mean absolute value (MAV) of a signal is the mean of the summation of absolute value and it is used in the characterization of bio-signal [160]. For $\mathbf{x}$, it can be defined as:

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|x_i|\,, \tag{5.8}$$

5.1.3.7. Modified mean absolute value

The modified mean absolute value ($MMAV$) is an extension of the MAV method. In $MMAV$, the signal is weighted by the window and a window is defined by two discrete values. For $\mathbf{x}$, it can be defined as [161]–[163]:

$$MMAV = \frac{1}{N}\sum_{i=1}^{N} w_i\,|x_i|\,, w_i = \begin{cases} 1, \text{ if } 0.25N \leqslant i \leqslant 0.75N \\ 0.5, \text{ otherwise} \end{cases} \tag{5.9}$$

where w_i is the weighting window.

5.1.3.8. Modified mean absolute value 2

The modified mean absolute Value 2 ($MMAV2$) is another extension of mean absolute value. In $MMAV2$, the signal is weighted by window function prior calculating the MAV [159]. $MMAV2$ is expressed using $\mathbf{x}$ as:

$$MMAV2 = \frac{1}{N}\sum_{i=1}^{N} w_i\,|x_i|\,, w_i = \begin{cases} 1, \text{ if } 0.25N \leqslant i \leqslant 0.75N \\ 4i/N, \text{ if } i < 0.25N \\ 4(i-N)/N, \text{ otherwise} \end{cases} \tag{5.10}$$

where, w_i is the weighting window function.

5.1.3.9. Log detector

The log detector (LD) is another feature extraction method which was also used to extract feature in biomedical signal processing [164]. Based on the logarithm and log detector (LD) feature, the nonlinear detector can be characterized a $log\left(|x_i|\right)$. For x, it can be expressed as [164]:

$$LD = exp\left(\frac{1}{N}\sum_{i=1}^{N} log\left(|x_i|\right)\right), \tag{5.11}$$

5.1.3.10. Permutation entropy

The permutation entropy (PE) is a simple and robust method. Through estimating the complexity of time series, it was used to identify the epileptic seizure [104], [165]. For a given time series x, each vector with d-th subsequent values is identified as:

$$i \mapsto \left(x_i, x_{i+1}, \ldots, x_{i+(d-1)}\right) \tag{5.12}$$

where d is the embedding dimension. An ordinal pattern linked with this vector identified as permutation $\pi = (k_0 k_1 \ldots k_{d-1})$ of $(01 \ldots d-1)$, which satisfies $x_{i+k_0} \leq x_{i+k_1} \leq \cdots \leq x_{i+k_{d-1}}$.

By considering a time lag τ eq. (5.12) can be further extended as:

$$i \mapsto \left(x_i, x_{i+\tau}, \ldots, x_{i+(d-1)\tau}\right) \tag{5.13}$$

For each time series, there is a probability distribution π, whose elements

Π_j ($j = 1, 2, \ldots, d!$) are the frequencies associated with the j possible permutation patterns. The PE can be defined as:

$$PE = -\sum_{j=1}^{d!} \Pi_j \log_2 \Pi_j, \tag{5.14}$$

In this study, the embedding dimension d and time lag τ were set to 3 and 1 respectively.

5.1.3.11. Spectral entropy

The spectral entropies are used to measure the complexity of a time series based on the power spectrum [105]. Several epilepsy seizure related studies [105]–[107] have proposed the use of spectral entropy such as Shannon (ShE) and Reny's entropy (RE). For a time series **x**, ShE and RE can be defined [105] as:

$$ShE = -\sum_{f} p_f \ln(p_f), \tag{5.15}$$

$$RE\,(\alpha) = \frac{1}{1-\alpha} \sum_{f} \ln p_f^2, \tag{5.16}$$

where α is the order of RE's entropy ($\alpha = 2$).

In eq. 5.15 and eq. 5.16, the normalization power p_f was calculated using Fourier transform of the time series x and it can be defined as :

$$p_f = \frac{P_f}{\sum P_f}, \tag{5.17}$$

where P_f is the power level of the frequency component.

5.1.4 Feature concatenation

For **x**, we calculate each feature using the above feature-extraction method from each segment and concatenate them sequentially expressed as a vector form. Therefore, the feature vector $\mathbf{v}_n$ of n-th subband for a channel can be defined as:

$$\mathbf{v}_n = [u_n^{(1)}, u_n^{(2)}, \ldots, u_n^{(D)}] \in \mathbb{R}^D, \tag{5.18}$$

where D indicates the number of total features (in our case D =12).

5.1.5 Subband and feature selection method

To select the relevant features and subbands, mutual information (MI) was used to estimate scores from the training set. Let us denote training features with the $M_n \in \mathbb{R}^{H\times D}$, where $H = ch \times s$ such that ch and s are the total number of channels and segments, respectively. The training features $M_n \in \mathbb{R}^{H\times D}$ were estimated for all channels with each segment and finally stacked all of the segments. The mutual Information (MI) from the set of the training features M_n and class C_n for n-th subband is defined sequentially as [166]:

$$MI(M_n; C) = \sum_{m_n^{(i)} \in M_n} \sum_{c \in C} p(m_n^{(i)}, c) \log \left(\frac{p(m_n^{(i)}, c)}{p(m_n^{(i)}) p(c)} \right), \tag{5.19}$$

where $m_n^{(i)}$ is the i-th feature and n-th subband. $p(m_n^{(i)}, c)$ is the joint probability of $m_n^{(i)}$ and c. The $p(m_n^{(i)})$ and $p(c)$ is the marginal probability density function of $m_n^{(i)}$ and c, respectively. We used bin-method [166] to estimate MI score between the features $m_n^{(i)}$ and label c from each feature of the training set M_n. The MI scores s_n from the training set M_n of n-th subband can be defined as:

$$s_n = [s_n^{(1)}, s_n^{(2)}, \ldots, s_n^{(D)}] \in \mathbb{R}^D, \tag{5.20}$$

Subbands scoring

From the eq. (14), we can calculate the MI scores of n-th subband as:

$$\bar{s}_n = \frac{1}{D}\sum_{j=1}^{D} s_n^{(j)}, \tag{5.21}$$

The set of mutual information, $\bar{s}_n$ for all subbands was rearranged in descending order such that $\bar{s}_{\lambda(1)} \geq \ldots \geq \bar{s}_{\lambda(n)} \geq \ldots \geq \bar{s}_{\lambda(L)}$, where $\lambda(n)$ is the sorted index of n subbands. The set of sorted MI scoring for all subbands can be defined as:

$$\mathbf{S}_{scores} = [\bar{s}_{\lambda(1)}, \bar{s}_{\lambda(2)}, \ldots, \bar{s}_{\lambda(L)}] \in \mathbb{R}^L, \tag{5.22}$$

Features scoring

The scores of d-th feature can be defined as:

$$\bar{s}^{(d)} = \frac{1}{L}\sum_{i=1}^{L} s_i^{(d)} \tag{5.23}$$

The set of average mutual information, $\bar{s}_d$ for all features was rearranged in descending order such that $\bar{s}^{I(1)} \geq \ldots \geq \bar{s}^{I(d)} \geq \ldots \geq \bar{s}^{I(D)}$, where $I(d)$ is the sorted index of d feature. The set of sorted average MI feature scoring for D features can be defined as:

$$\mathbf{F}_{scores} = [\bar{s}^{I(1)}, \bar{s}^{I(2)}, \ldots, \bar{s}^{I(D)}] \in \mathbb{R}^D. \tag{5.24}$$

Finally, we apply a grid-search method between the subbands and features with the higher values of scores ($\mathbf{S}_{scores}$ and $\mathbf{F}_{scores}$) and F-score was estimated. To select the significant features and subbands the value with

maximum F-score value.

$$\mathbf{V}^* = [\tilde{\mathbf{v}}_{\lambda(1)}, \ldots, \tilde{\mathbf{v}}_{\lambda(n)}, \ldots, \tilde{\mathbf{v}}_{\lambda(G)}] \in \mathbb{R}^{G \times H}, \tag{5.25}$$

where $\tilde{\mathbf{v}}_{\lambda(n)} = [u_n^{I(1)}, u_n^{I(2)}, \ldots, u_n^{I(H)}]$ for the n-th subbands.

5.1.6 Classifiers

Support vector machine

The SVM uses an optimal hyperplane to separate data from two classes [167]. To operate nonlinear relationship between dependent and independent variables, the RBF kernel maps data into a higher-dimensional space in nonlinear way. Based on the training set, the parameter of an SVM was set. In this chapter, we used an SVM with a radial basis function kernel (RBF). It searches all the current leaves each time to find the leaf with the largest splitting gain.

LightGBM

LightGBM uses the gradient boosting decision tree algorithm for classification used in machine learning for epilepsy seizure-detection [168]. Traditional boosting method uses the level-wise decision tree growth strategy that splits all leaves every time whereas LightGBM uses the leaf-wise strategy with depth limitation. For splitting, each time it searches all the current leaves that has largest splitting gain. So, LightGBM can effectively reduce the calculation time, reduce the error as well as improve the classification accuracy compared to traditional boosting methods. The detailed derivation and additional information are available in [169]. In study, we used 5-min iEEGs data were used to optimize the parameters of iteration trees.

5.1.7 Evaluation

5.1.8 Division of iEEG time-series for training and testing

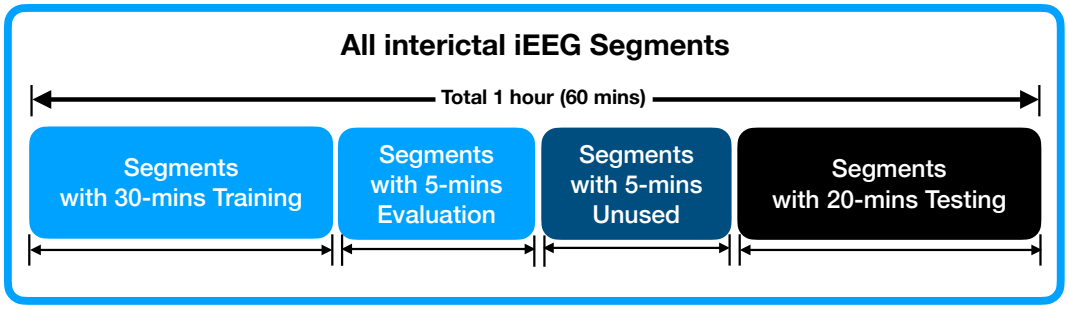

FIGURE 5.2: The splitting the iEEG data used for training and testing in the proposed method.

To design a mathodological framework for detecting the SOZ and non-SOZ channels, we have divided the data into training, validation, and testing set. By considering the nature of the time series, we have used time series cross-validation techniques [170]–[172], one of the suitable solutions for model evaluation. In this study, we used 1 hour interictal iEEG data and divided them into 20-s segments. Among 180 segments, 90 segment (30-min iEEG data) used for training, 15 segments (5 mins) used for tuning the parameters. Other remaining 75 segments (25 min iEEG data), 15 segments (5 mins) was unused and 60 segments (20 mins) used for testing. Fig. 5.2 shows the splitting process of 1 hour interictal iEEG data of our proposed design. Due to the imbalanced of SOZ and non-SOZ channels in iEEG data, the number of non-focal segments in non-SOZ channels are much higher than the number of focal segments in SOZ channels. Several studies have observed the problems of using machine learning methods for the imbalanced distribution in minority and majority classes [119]–[121]. Our previous study for epileptic focus detection has provided an evidence that the imbalanced learning problem can deteriorate the performance of the design [53]. To solve the imbalanced learning problem, the adaptive synthetic (ADASYN) approach

[120] was used. Let us defined the training features $\nu = \left\{\mathbf{v}_F^{*(i_{in})}, \mathbf{v}_{NF}^{*(j_{in})}\right\}$ after selecting subbands and features indeced from eq. (22), where $\mathbf{v}_F^{*(i_{in})}$ and $\mathbf{v}_{NF}^{*(j_{in})}$ denote the feature vector of the i_{in}-the sample of focal segement and the feature vector of the j_{in}-th sample of non-focal segment, respectively and $I_{in} \ll J_{in}$ due to imbalanced dataset. The balance training set $\tilde{\nu}$ was defined in [120] from the training set ν as:

$$\tilde{\nu} = \left\{\mathbf{v}_F^{*(i_{in})}, \mathbf{v}_{NF}^{*(j_{in})}, \mathbf{v}_F^{*(\tilde{i}_{in})}\right\}, \tag{5.26}$$

where, $I_{in} + \tilde{I}_{in} = J_{in}$. In this study, the parameters to balance the training set were set to the similar to our previous study [53]. The balance training set $\tilde{\nu}$ was the input of SVM method for training the model.

Segment wise performance measurement

In this study, the performance evaluation metrics include sensitivity (SEN), specificity, false discovery rate (FDR), and F-score (F_1 score) to measure the performance of segment-wise detection in SOZ and non-SOZ channels. The underlying idea behind showing statistics of segment-wise detection was that we provided measures for comparison studies similar to HFO-related works [51], [149] and low frequency-based studies used in Bern-Barcelona or other datasets. These evaluation metrics also used in the recent HFC-related based research [53]. The calculations of evaluation metrics are as follows:

$$Sensitivity = \frac{TP}{TP + FN} \times 100\%, \tag{5.27}$$

$$Specificity = \frac{TN}{TN + FP} \times 100\%, \tag{5.28}$$

$$FDR = \frac{FP}{TN + FP} \times 100\% \tag{5.29}$$

$$F_1\ score = \frac{TP}{TP + 0.5(FP + FN)} \tag{5.30}$$

where true positive (TP) denotes the number of correctly detected SOZ segments in the SOZ channels; false negative (FN) refers to the number of incorrectly detected non-SOZ segments in the SOZ channels. true negative (TN) indicates the number of correctly detected non-SOZ segments in non-SOZ channels; false positive (FP) means the number of incorrectly detected SOZ segments in the non-SOZ channels. A post-hoc test with the Bonferroni procedure was used to assess the statistical significance of the methods with a significance level of $\alpha(= 0.05)$

Channel wise performance measurement

In this study, the main target was to design the automatic system to identify the possible electrodes related to SOZ. To identify the electrodes with SOZ and non-SOZ, the final decision will come after observing the scores of multiple segments. Therefore, the performance of each patient was observed by AUC-ROC [128] by computing the sensitivity and FPR of the channels with each threshold values. In our case, we estimated scores from each segment of test set and averaged togather to achieve the final score of the channels. After achieving SEN and FPR of the channels, we estimated the AUC by using the trapezoid rule [128].

5.2 Experimental results

In this study, we have developed our framework based on a different cases. For each of these cases, we have investigated the improvement of the design performance with statistical evidence. First, we have discussed the cases for both SVM and LightGBM classifier. Then we discussed the optimal features and subbands selection in Sec. 5.2.1. After selecting the optimal features and subband, we have provided an intuition with statistical measurements based on optimal method in Sec. 5.2.2. In Sec. 5.2.3, we give the localization result of SOZ and non-SOZ channels. To evaluate the possibility of patient-independent design (PID), we compared the proposed design with PID on data from eleven patients shown in Sec. 5.2.4. Finally shows the analysis of computational cost in Sec. 5.2.5. The cases that are considered in the proposed detection design for both classifier (SVM and LightGBM) are summarized below:

Filter-bank feature extraction method (FbFM): The multichannel interictal iEEG signals are splitted into 20-s segments. The N bandpass filters were implemented using a third-order Butterworth filter to decompose each segment of the high-frequency components (ripple and fast ripple) in iEEG. The different types of statistical feature extraction methods are applied onto each subband to extract features. The SVM and LightGBM with ADASYN method are used to score each electrode for identifying possible SOZ channels.

FbFM with subband and feature selection (FBFM/Sb/FS): In this case, subbanding and feature extraction were performed in the same way as the above **(FbFM)** method. A data-driven grid-search method using MI scores are proposed to select both prominent bands and features. The ADASYN approach, SVM and LightGBM classifier are also used to score of channels.

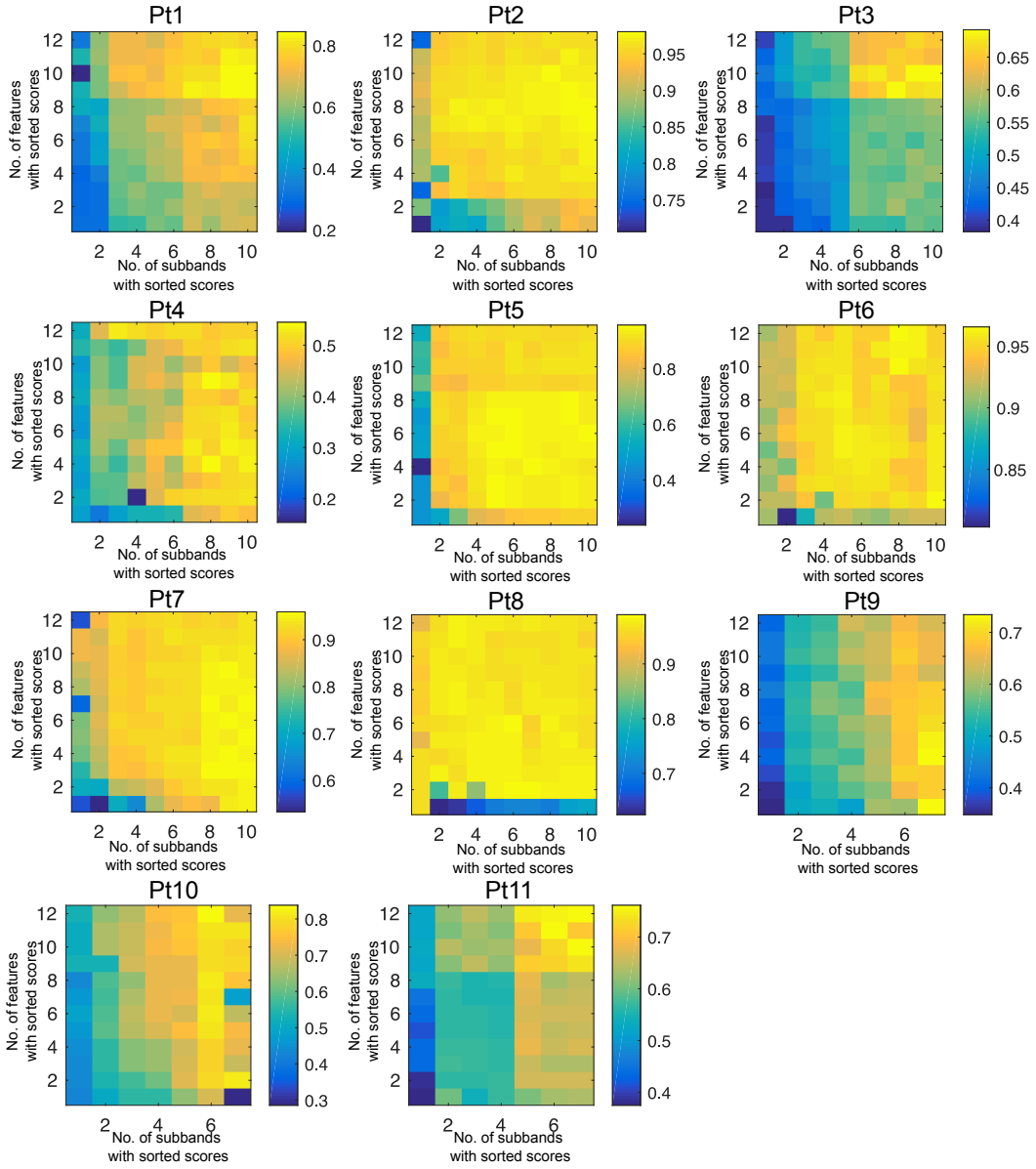

FIGURE 5.3: Grid search method for selecting prominent features and subbands. The F-score was acquired as linked combination of subbands (X-axis) and features (Y-axis) with highest MI scores.

5.2.1 Selection of optimal features and subbands

Different high frequency oscillation (HFOs) related studies [48]–[50] suggested that to identify the SOZ area only need to observe ripple and fast ripple bands. But, to find the SOZ through the observation of wider bands in the high-frequency is difficult task for the epileptologist. Recently, EEG based BCI study [173] have shown that, the selection of appropriate operational

bands from wider band can significantly improve the performance of the system. However, selection of optimal band with prominent features jointly are a difficult task. But considering this problems, in this study we have developed a data-driven grid-search method based on mutual information scores to select the prominent features and epileptic subbands with joint contributions, improving the performance of design framework. The proposed method computes MI scores for the each of the features and subbands. The subbands and features are rearranged according to their maximum values of mutual information scores. A grid search method was applied to select optimal features H and subbands G by aggregating the top ranked MI scores of subbands and features.

Fig. 5.3 shows the selection of optimum features and subbands with respect to F-score for all eleven patients. This figure shows the relationship between the F-score and the number of incorporated subbands and features with the higher value of MI scores ($\mathbf{S}_{scores}$ and $\mathbf{F}_{scores}$ induced from eq. 5.22 and eq. 5.24., that may lead to improving the performance of our proposed design.

5.2.2 Results for detected segments

In this study, we proposed the methods such as FbFM, and FbFM/Sb/FS with respect to two classifier (SVM and state-of-the-art LightGBM) to detect the SOZ and non-SOZ channels of eleven patients. The performance of different methods in term of sensitivity, specificity, FDR, and F_1 score is a widely-used metric to evaluate the system for imbalanced dataset. Table 5.2 shows the experimental results for individual segment detection for SVM classifier. From this table it has been observed that method FbFM/Sb/FS provide improve performance compared to FbFM methods. Similarly, the

segment detection result for LGBM classifier shown in Table 5.3. The sensitivity represents the detection of correctly predicted focal segments from the SOZ channels (Pt1: 82.78%; Pt2: 99.83%; Pt3: 67.77%; Pt4: 81.11%; Pt5: 90.83%; Pt6: 90.24%; Pt7: 83.83.00%; P8: 99.69%; P9: 89.33%; P10: 76.83%; and P11: 80.33%). In contrast, the FDR is a evaluation metric used to characterize the rate of incorrectly prediction focal segments from the non-SOZ channels (Pt1: 0.56%; Pt2: 1.08%; Pt3: 9.91%; Pt4: 6.52 %; Pt5: 1.79%; Pt6: 0.23%; Pt7: 1.22%; P8: 0.31%; P9: 1.79%; P10: 3.67%; and P11: 0.33%). To test the statistical significance of the methods (FbFM vs FbFM/bS/FS), the result of the post-hoc tests with F_1 score were performed to observe the significancy of the methods performance. From the results of post-hoc tests, the FbFM/bS/FS method with LightGBM significantly outperformed the FbFM method (FbFM/Sb/FS vs FbFM with SVM: $p < 0.05$; FbFM/bS/FS vs FbFM with LightGBM: $p < 0.05$; FbFM/Sb/FS with LightGBM vs FbFM/Sb/FS with SVM: $p < 0.05$). A similar scenario is observed for the other evaluation metrics. Consistently, the proposed FbFM/Sb/FS approach with LightGBM achieves the highest performance for all patients. Considering the overall results, the method FbFM/Sb/FS with LightGBM are used as an optimal method for futher analysis in this study.

5.2.3 Results for localization of SOZ channels

The previous section such as Sec. 5.2.2 already confirmed that the joint contribution of selected bands and features can significantly improve the performance of proposed method. The performances of the optimal method FbFM/Sb/FS with LightGBM for identifying SOZ were observed in terms of AUC shown in Table 5.4.

The Fig. 5.4 and Fig. 5.5 illustrated the visualization of channels with MRI

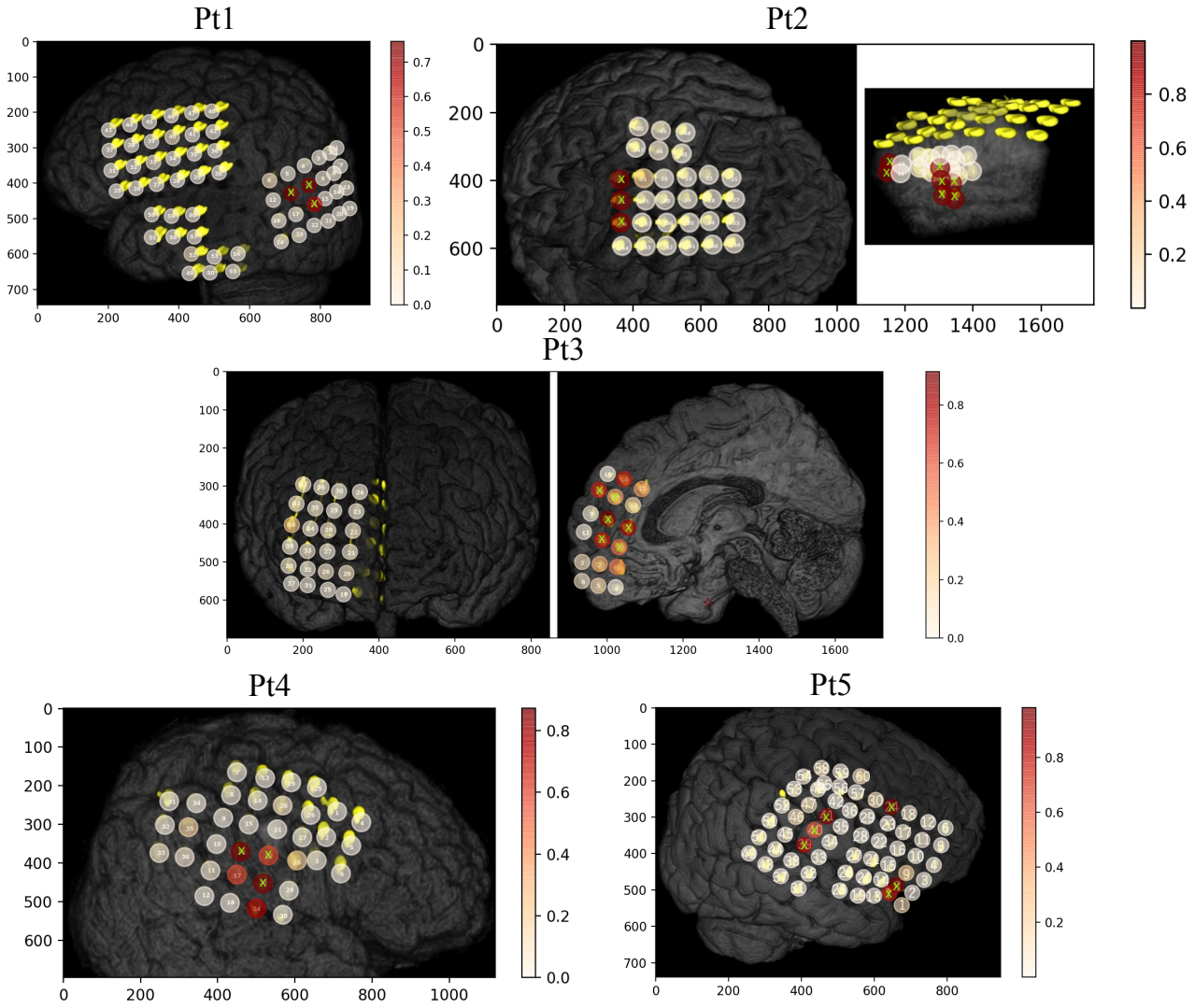

FIGURE 5.4: The simulated results using the FbFM/Sb/FS method with MRI image for Pt1 to Pt6. For each patient, the "X" with lime color indicates the SOZ electrodes labeled by clinical expertise and the circles with color represents the average scores of the channels estimated by our proposed method.

scan images in xy-plane. The "X" with lime color indicates the SOZ electrodes labeled by clinical expertise and the circles with color represents the scores of the electrodes estimated by our proposed method. For each channel, the estimated score values were plotted onto the area of cortical surface which provides a graphical view to the epileptologist about the detected SOZ and non-SOZ electrodes. The scores of each channel are achieved by averaging the scores across all of the segments. From Fig. 5.4 and Fig. 5.5, it is observed that, our have proposed computer aided design provides higher

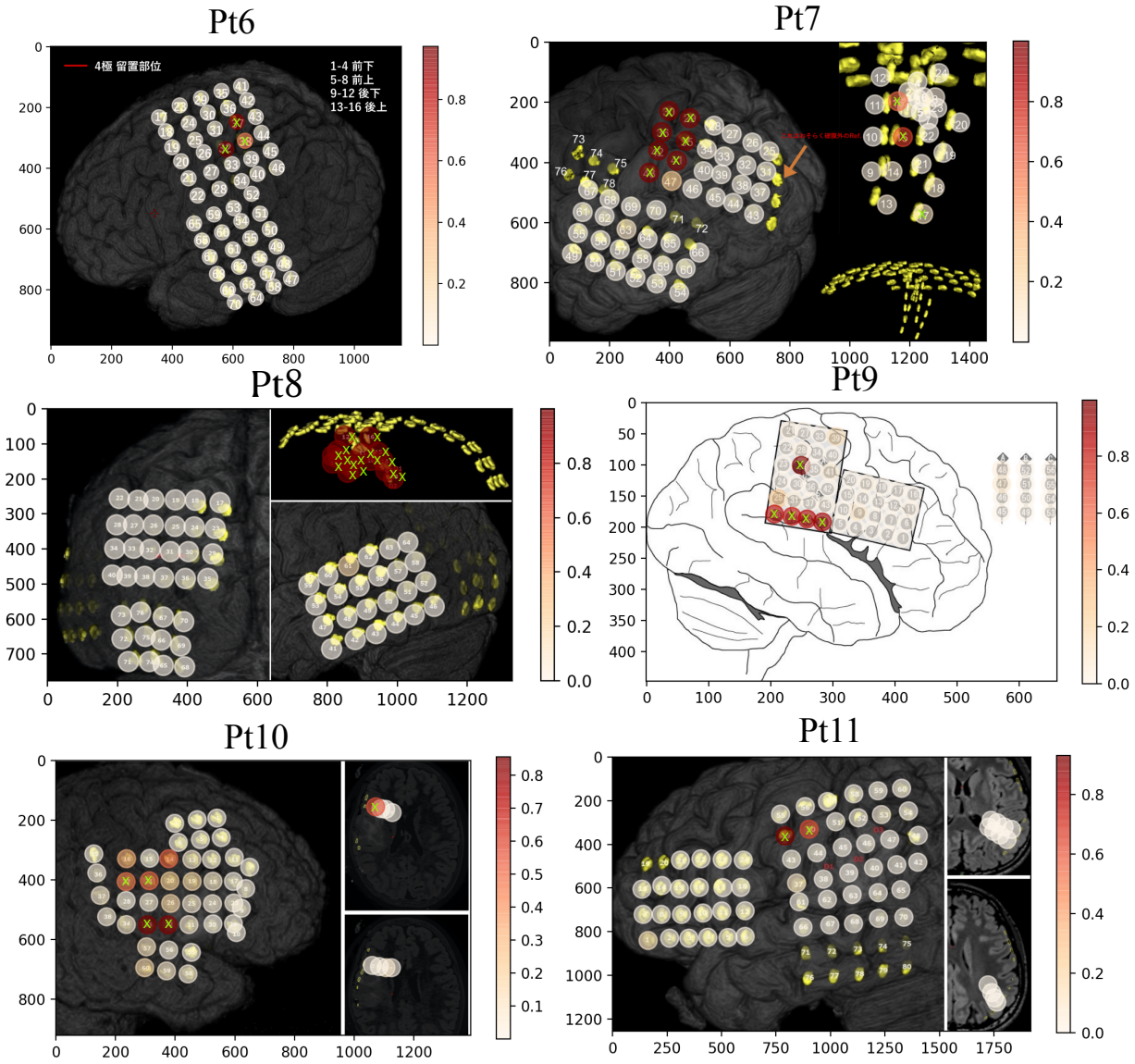

FIGURE 5.5: The simulated results using the FbFM/Sb/FS method with MRI image for Pt7 to Pt11.6. For each patient, the "X" with lime color indicates the SOZ electrodes labeled by clinical expertise and the circles with color represents the average scores of the channels estimated by our proposed method.

scores to that electrodes labeled by clinical experts (red color with circle in the Fig. 5.4 and Fig. 5.5) and also suggests some active electrodes that are close to the SOZ and may have linked to seizure event.

5.2.4 Comparison between patient-dependent and -independent designs

In fact, for the real world application patient-independent design (PID) is more preferable compared to patient-dependent design (PDD). Therefore, it is desirable to compare the proposed PDD in order to assess the possibility of designing PID. Thus, in this study, we compared the proposed PDD and PID with AUC for localizing SOZ-channels using optimal classifier Light-GBM. Then, the resulting model was tested using given data of test subject for testing the model. To balance the class features, the ADASYN with default setting [53], [120] is also applied to highly imbalanced feature sets in the training stage. The obtained results are displayed in Table 5.4 for both PDD and PID. As mention that, for patients with 1 kHz sample frequency (see in Table 5.1) was between 100 Hz to 450 Hz. Hence, to design the PID framework, we have used only the bands between 100 Hz to 450 Hz for patients with 2 kHz sampling frequency. To develop patient independent method for patients with 2 kHz sampling frequency, we have eradicated the patients from the stacked training data with 1 kHz sampling frequency.

5.2.5 Computational time analysis

The mean computational time for each feature extraction method (12 methods) using 10 subbands was measured by Python on iMac Pro (with Intel Xeon W processor and 128 GB RAM). Of note, to detect a single segment at the testing phase, the mean results are estimated with 100 runs. Fig. 5.6 shows the bar diagram (left) of the average computational time (in seconds) with each feature extraction methods for ten subbands. From this figure, it is observed that, the computational time of some methods (MAV, RMS, and ShE, RE) are less than 0.02-s. So, for a single segment test, the average computational time with the combination of twelve feature-extraction methods

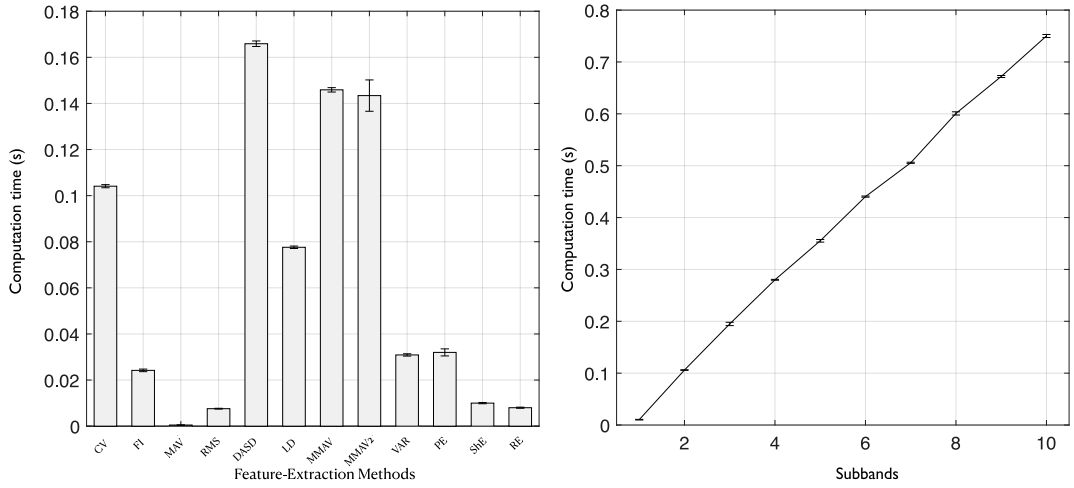

FIGURE 5.6: Mean computational time with twelve statistical features including information theoretic features for ten sub-bands (left). Average computational time with number of sub-bands for twelve feature-extraction methods (right).

with ten sub-bands are 0.75-s. Fig. 5.6 also shows the mean computational time (in seconds) with increase number of sub-bands for twelve feature extraction methods (right). It is observed that the computational time is increased linear fashion with increasing number of subbands. Since the total computational time for a single segment test is 0.75-s, it is very convenient for clinical application and epileptologists can take quick medical decision from long term iEEG data.

5.3 Discussion

Apart from biomarker, in this study, we have used high-frequency components (>80 Hz) such as ripple and fast ripple bands for detecting the SOZ electrodes. To identify the SOZ electrodes, in a conventional clinical system, the epileptologists visually inspect more than 3 days long iEEG data (depends on patient's condition) to observe the extent of interictal epileptic discharges (IEDs) and seizure discharges. The epileptologists decide (or diagnose) SOZ channels in order to carry out epilepsy surgery via an epileptic focal resection. In this study, we proposed a method for identifying SOZ

electrodes in HFCs (>80 Hz) of iEEGs with selection of prominent bands and features. In our methodology, firstly, we detected the individual segments by using SVM and LightGBM with different conditions (FbFM and proposed FBFM/Sb/FS). We attained compatible performances with only 35 mins of interictal iEEG signal that make the proposed FbFM/Sb/FS method with LightGBM classifier more interesting to design SOZ detection. Then, we have simulated the results to determine the possible SOZ electrodes, that may assist epileptologists to hypothesize about the SOZ and non-SOZ electrodes. In our analyzed dataset described in Table 5.1, among eleven patients three patients (Pt5, Pt9, and Pt10) are not seizure free (”residual”). Although for ”residual” patients, the SOZ electrodes labeled by clinical experts were removed by surgery. However, for residual patients, our proposed methodology gave suggestions for some active electrodes, which were very close to the SOZ area. This localization resutt allows clinical expert to make a more confident decision with their expertise.

Besides, for possibility assessment of the patient dependent design (PID), we have calculated the AUC and compared with the patient-dependent design (PDD). We observed that, compared to our proposed PDD method, the result of PID is very poor. One of the possible reason may be the subject-dependent nature of EEG signals and very different locations of electrodes. However, the major advantage of the patient independent design is that, to detect SOZ electrodes for new patients do not need any labeled data. Varatharajah et al. [42] proposed an artificial intelligence based framework for identifying the SOZ electrodes using SVM classifier. In their study, they used electrophysiological biomarkers including HFO, IED, and phase-amplitude coupling (PAC), between low-frequency bands (0.1–30 Hz) and high-frequency bands (65–115 Hz). To evaluate the system, they used 82 patients with medically intractable epilepsy. For mixed data framework, they mixed the SOZ and NSOZ (non SOZ) electrodes of all subjects to form a dataset and their

acquired average AUC was of 0.79. In patient independent design, they used leave-one-patient-out cross validation with average AUC of 0.73. The major shortcoming of their proposed framework was that, they only considered the partial ripple bands. In our PID framework, we have considered both ripple and fast ripple bands and our obtained average AUC of 0.68. Although, we have used only eleven patients which is much smaller compared to Varatharajah et al. study, but our PID result was very closest to their proposed work. However, in conventional clinical studies, HFO, IED, and PAC are deliberated as important biomarkers, which were not considered in our proposed study. In fact, from the machine learning viewpoint, the statistical features are significant to characterize the epileptic event. So, instead of biomarkers including HFO, IED, and PAC, our statistical features are very good candidates for identification of SOZ electrodes. Besides, there is still room for improvements of PID in terms of AUC. To improve the performance of the PID, future study could be more advanced signal processing method with increased number of patients. A more hopeful direction could be the usage of domain transfer to adapt the different data distributions [174]

For a direct comparison of other epilepsy-related studies to design an automatic system, we have summarized the low frequency related studied [39]–[42], [175] in Table 5.5. Similarly high frequency components (HFC) and high frequency oscillation (HFO's) related studies are summarized in Table 5.6. To identify the focal epilepsy, Sharma et al. [39] proposed a system using Bern-Bercelona dataset with 87% of accuracy. To enhance the performance of the system, the similar problem also introduced by different studies [40], [41], [43], [139]. In their proposed system, they have used different decomposition techniques such as wavelet, EMD etc. with variant of classifiers including Naïve Bayes (NBC), radial basis function (RBF), SVM, k-NN, non-nested generalized exemplars (NNge), and best first decision tree (BFDT). For identifying focal epilepsy, all of the above LFC related epilepsy studies (shown

in Table 1) used the Bern-Barcelona dataset that consisted of approximately 20 s of intracranial EEG. The dataset has a pair of focal and non-focal channels (3750 focal and 3750 non-focal) with frequency band was between 0.5 and 150 Hz. However, the limitations of all above discussed low frequency related studies was that, they have used low frequency components (0.5–150 Hz) with well balanced problem.

Several HFO-related epilepsy studies [51], [149], [150] have proposed the automatic HFOs detector that are shown in Table 5.6. In their proposed system, they hypothesized that the rate of HFOs tends to be higher in seizure onset zone. However, to calculate the baseline, most of the those HFO studies used the long-term intracranial EEG data and for ripple and fast ripple bands, they designed the automatic system individually. Recently, our published study used eight types of information theoretic entropy based features extraction methods for identifying the SOZ electrodes [53]. For methodological evaluation, we have used 10-fold cross validation with SVM classifier. For segment detection, an average sensitivity of 63.03%, and specificity of 91.85%. For identification of the channels, an average AUC of 0.88. However, our proposed entropy based studied has some limitations that already discussed in introduction section. To address those problems, in this study, we have proposed a methodological framework with considering 1) the division of the data into training, validation, and testing set with the way of time series forecasting; 2) the use of statistical feature-extraction methods to reduce the parameter selection problems as well as computational time; 3) the selection of most significant subbands, so that the neurologists may focus to the specific narrow bands to spot the epileptic symptoms. Our proposed PDD method achieved an average sensitivity of 85.73% and an average specificity of 97.50% for segment detection. For detection of the channels an average AUC of 0.99. Note that, the proposed method used the combination of 60 segments in the test phase to identify the channels.

Note that the proposed method used the combination of 60 segments in the test phase to identify the channels. The the average computational time with 12 feature-extraction methods to identify a segment was 0.75 s. Sheuli et al. reported that the average computational time of the methodological framework with eight entropies and 10 subbands was 56.51 s to test a single segment with 60 channels [53]. However, assuming that the testing data with 60 segments, the computation cost of the framework would be close to 60 min to identify SOZ and non-SOZ channels. The proposed method could facilitate this procedure (0.2-min) without compromising performance, which could improve the usability of the framework.

TABLE 5.2: Experiment results for individual segment-detection with eleven patients using FbFM and proposed FbFM/Sb/FS with SVM.

Patient ID	Evaluation matrics (FbFM with SVM)			
	Sensitivity	**Specificity**	**False discovery rate (FDR)**	**F-score**
Pt1	62.77	99.32	0.67	0.71
Pt2	84.00	94.87	5.12	0.82
Pt3	58.61	87.92	12.08	0.50
Pt4	34.44	**93.78**	**6.21**	0.33
Pt5	67.50	97.62	2.37	0.71
Pt6	90.48	94.47	5.52	0.75
Pt7	75.50	97.88	2.11	0.80
Pt8	92.60	99.53	0.47	0.95
Pt9	70.66	91.53	8.46	0.55
Pt10	47.66	91.80	8.20	0.51
Pt11	86.66	97.14	2.85	0.62
Average	70.08	95.08	4.92	0.67
Patient ID	**Evaluation matrics (FbFM/Sb/FS with SVM)**			
	Sensitivity	**Specificity**	**False discovery rate (FDR)**	**F-score**
Pt1	73.33	99.39	0.61	0.79
Pt2	99.39	96.42	3.58	0.93
Pt3	**81.94**	87.59	12.41	**0.63**
Pt4	77.78	93.08	6.91	0.61
Pt5	69.44	97.78	2.22	0.73
Pt6	**90.95**	96.24	3.75	0.80
Pt7	81.83	97.41	2.58	0.82
Pt8	97.81	99.72	0.27	0.98
Pt9	69.33	95.98	4.02	0.66
Pt10	53.16	95.43	4.57	0.60
Pt11	**89.16**	**98.36**	**1.64**	0.73
Average	80.42	96.13	3.88	0.76

TABLE 5.3: Experiment results for individual segment-detection with eleven patients using FbFM and proposed FbFM/Sb/FS with LightGBM.

Patient ID	Evaluation matrics (FbFM with LightGBM)			
	Sensitivity	**Specificity**	**False discovery rate (FDR)**	**F-score**
Pt1	78.33	98.62	1.37	0.76
Pt2	99.33	97.16	2.83	0.94
Pt3	67.50	85.60	14.39	0.53
Pt4	76.67	91.11	8.88	0.55
Pt5	86.11	97.59	2.41	0.82
Pt6	90.00	**99.80**	**0.20**	0.94
Pt7	83.33	97.16	2.83	0.83
Pt8	92.70	99.52	0.47	0.95
Pt9	88.66	93.36	6.63	0.69
Pt10	74.50	95.80	4.20	0.76
Pt11	83.33	99.06	0.93	0.77
Average	83.68	95.89	4.17	0.78
Patient ID	**Evaluation matrics (FbFM/Sb/FS with LightGBM)**			
	Sensitivity	**Specificity**	**False discovery rate (FDR)**	**F-score**
Pt1	**82.78**	**99.44**	**0.56**	**0.86**
Pt2	**99.83**	**98.92**	**1.08**	**0.98**
Pt3	67.77	**90.09**	**9.91**	0.60
Pt4	**81.11**	93.48	6.52	**0.64**
Pt5	**90.83**	**98.20**	**1.79**	**0.88**
Pt6	90.24	99.77	0.23	**0.95**
Pt7	**83.83**	**98.78**	**1.22**	**0.88**
Pt8	**99.69**	**99.69**	**0.31**	**0.99**
Pt9	**89.33**	**98.20**	**1.79**	**0.86**
Pt10	**76.83**	**96.33**	**3.67**	**0.79**
Pt11	80.33	99.67	0.33	**0.84**
Average	**85.73**	**97.50**	**2.49**	**0.84**

TABLE 5.4: The results of AUC to detect the SOZ channels for both patient-dependent and -independent design.

Patient ID	AUC	
	Dependent	Independent
Pt1	1.00	0.77
Pt2	1.00	0.72
Pt3	0.97	0.64
Pt4	0.98	0.55
Pt5	1.00	0.66
Pt6	1.00	0.63
Pt7	0.99	0.77
Pt8	1.00	0.65
Pt9	1.00	0.58
Pt10	0.99	0.55
Pt11	1.00	0.98
Mean	**0.99**	0.68

TABLE 5.5: Low frequency component related comparative studies

Low frequency components related studies					
Reference.	Sharma et al. [39]	Itakura et al. [40]	Arunkumar N et al. [41]	Varatharajah et al. [42]	Yang Y. et al. [175]
Dataset	Bern-Barcelona	Bern-Barcelona	Bern-Barcelona	Mayo Clinic	Bern-Barcelona
Methods	-EMD 6-entropy based feature -SVM	-BEMD -6-entropy based feature -LS-SVM	-3-entropy based features -NBC, SVM, k-NN, RBF NNge, BFDT	-PAC, HFOs, IEDs -SVM	-FAWT -2-entropy based feature -GRNN, SVM, LS-SVM, k-NN fKNN
Bands	Lower bands (0.5–150 Hz)	Lower bands (0.5–150 Hz)	Lower bands (0.5–150 Hz)	Lower bands (0.1–30 Hz) Partial ripple bands (65–115 Hz)	Lower bands (0.5–150 Hz)
Goal	Epileptic focus detection	Epileptic focus detection	Epileptic focus detection	SOZ detection	Epileptic focus detection
Performance	ACC: 87%	ACC: 86.89%	ACC: 98.0%; Sen: 100%; Spe: 96.0%	For cross-validation of mixed data AUC: 0.79 and For leave one patient out cross-validation AUC: 0.73	ACC: 94.80%

TABLE 5.6: High frequency oscillation (HFOs) and High frequency component (HFC) related comparative studies

High frequency oscillation (HFOs) related studies			
Reference	Jrad et al. [149]	Zuo et al. [51]	Lai et al. [150]
Dataset	Rennes University Hospital	Xuanwu Hospital	West China Hospital
Methods	-Gabor transformation -Energy-based feature -SVM	-Deep CNN -Time-frequency map	-CNN method
Bands	Ripple (120–250 Hz) Fast ripple (250–600 Hz)	R (80–250 Hz) FR (250–500 Hz)	R (80–250 Hz) FR (250–500 Hz)
Goal	HFOs detection	HFOs detection	HFOs detection
Performance	Ripple-HFOs (Sen: 81.1% and FDR: 30.2%) and FR-HFOs (Sen: 74.6% and FDR: 6.3%)	R-HFOs (Sen: 77.0% and SPE: 72.3%) and FR-HFOs (Sen: 83.2% and Spe: 79.3%)	R-HFOs (Sen: 82.2% and FDR: 12.6%) and FR-HFOs (Sen: 93.4% and FDR: 8.0%)

High Frequency Components (HFC) related studies		
Reference	Akter et al. [53]	Proposed
Dataset	Juntendo Hospital	Juntendo Hospital
Methods	-Multiband -Eight type of entropy -Feature selection -SVM	– Multiband –12 type of features -Joint bands -Feature selection and -LightGBM
Bands	R and FR bands (100–600 Hz) 8-patients 2 kHz	R and FR bands (100–600 Hz) and (100–450 Hz) 11-patients 2 kHz and 1 kHz
Goal	SOZ detection	SOZ detection
Performance	Segment detection (Sen: 63.03%; Spe: 91.85%) and SOZ detection (AUC=0.88) Com. time: 56.51 s	Segment detection (**Sen: 85.73**%; **Spe: 97.50**%) and SOZ detection (**AUC=0.99**) Com. time: **0.75s**

Chapter 6

Conclusion and Future Work

6.1 Conclusions

In this chapter, we have concluded the study. In our research, we have developed an artificial intelligence based computer aided solution for identification of seizure onset zone (SOZ) electrodes from high-frequency components including ripple and fast ripple of interictal intracranial EEG data. The contributions of our proposed method are as follows:

6.1.1 Graphical representation for visualization

We have first time introduced a graphical representation of our proposed SOZ detection result, that shown in Chapter 4. This simulated results (see in Fig. 6.1), represent the localization of SOZ and non-SOZ channel which help the epileptologists in two ways: (1) to observe the localization of SOZ and non-SOZ segments over duration of the iEEG data, and (2) the number of detected segments corresponding to the SOZ and non-SOZ channels.

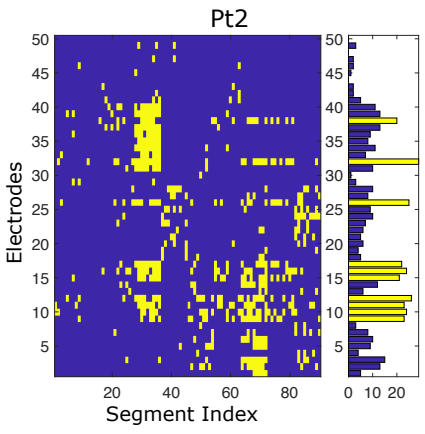

FIGURE 6.1: Graphical representation with segment wise detection for visualization

6.1.2 Provides suggestion about the active electrodes closest to the seizure onset zone

In this thesis, we have provided the simulated result with MRI scan image like Fig. 6.2. This intuition can gives some useful suggestions about the active electrodes, which are very closest to the SOZ electrodes. Moreover, it can provides the illustration to hypothesize the the possible SOZ channels much easier and more reliable way. Besides, it can assists the epileptologist to take quick medical decision based on scores of the channels mapped into the cortical surface using MRI images.

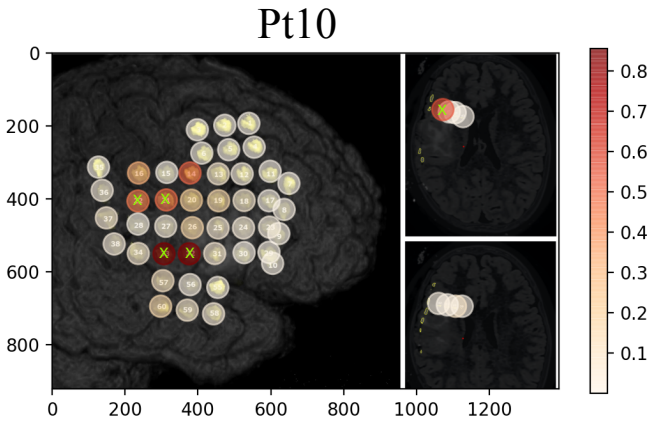

FIGURE 6.2: Visualization of electrodes with MRI scan image.

6.1.3 Used high frequency components (HFCs) for localization of SOZ electrodes

It is known that the activity in the high frequency components, including ripple and fast ripple bands, of interictal iEEG, are associated to the epileptic seizure. For identifying the SOZ electrodes, instead of using high frequency oscillations which need a long term interictal iEEG signal to calculate the baseline, we have used high frequency components such as ripple and fast ripple bands. In this thesis, we have used only 30 minutes (Chapter 4) and one hour (Chapter 5) of interictal iEEG signal with high frequency components for detecting the SOZ electrodes.

6.1.4 Introduced efficiently work statistical features for SOZ localization

In this study, we have proposed twelve statistical feature extraction methods including information theoretic features, which are important to characterized the epileptic events. Although, in epilepsy related studies some biomarkers such as HFO, IED, and PAC are considered very important in conventional clinical studies. But, our proposed study does not use biomarkers such as HFO, IED, and PAC. In fact, from the machine learning point of view, the statistical features are very significant to characterized the epileptic events. In this thesis, we have achieved very promising result for all patients by using these statistical features with LightGBM classifier shown in Chapter 5. Hence our statistical features are very good candidate instead of different biomarkers.

6.1.5 Jointly selection of features and subbands

In this thesis, a data-driven grid-search method using mutual information scores is developed to optimize subbands and features jointly. The joint selection of appropriate bands and features that are related to epileptic activities improved the performance of our proposed computer-aided solution and it is our first proposed method for SOZ detection related epilepsy studies. The jointly selection of optimal features and subbands shown in Chapter 5. Such joint selection can assist the epileptologist during data analysis and may reduce the workload of them.

6.1.6 Used real world clinical data

In this work, we have evaluated the proposed method for eight epilepsy patients (Chapter 4) and eleven patients (Chapter 5) considering different ages (adult and pediatric) and pathological types. We have used real world long

term interictal iEEG data collected from Juntendo University Hospital, which was approved by the ethics committee of Juntendo University Hospital as well as the Tokyo University of Agriculture and Technology, Japan.

6.1.7 Used focal cortical dysplasia (FCD)

In this work, we have used FCD because, it is the most common reason of medically intractable epilepsy in both pediatric and adult patients. Since, the objective of this study was to develop a machine learning methodology for identification of the seizure onset zone (SOZ), the type of epilepsy for all patients is preferred to be identical for evaluation of the methodological framework. Moreover in FCD, it is relatively easy to localize the SOZ electrodes because its labels reliability is high in machine learning algorithms.

6.2 Future work

6.2.1 To implement a patient-independent design

In this thesis, we have proposed a methodology for patient-dependent design (PID) for SOZ identification. We have used only 30-mins of signal of interictal phase in the proposed priori-based system and the epileptologists need to label the electrodes with only 30-minutes of the interictal iEEG for new patient to identify the possible SOZ electrodes. In addition, to reduce the workload of epileptologist, in this thesis we have also considered the patient independent case. The advantage of patient-independent design (PID) is that, we do not need any labeled data for the patients to detect the SOZ area. However, in PID case, the result is very poor compared to the patient-dependent design (PDD). One of the possible region of this result will be the subject-specific nature of iEEG signals. In machine-learning research, several studies proposed the use of domain transfer learning to adapt the different

distributions of features extracted from different subjects [174], [176]. To enhance the performance of the PID case, future works could study more advanced signal processing methods. A more hopeful direction could be the practice of domain transfer to adjust the different distributions [174].

6.2.2 To use other pathology in future

In this study, we have used only FCD pathology type data. But for identification of SOZ through analysis of HFO, other pathology types are also used including dense gliosis (DG), hippocampal sclerosis (HS), low grade glioma (LGG) etc. As a future work, we will try to use another pathology diagnosis for identification of SOZ electrodes.

Appendix A

List of Publications

Journal Papers

1. **Most. Sheuli Akter**, Md. Rabiul Islam Toshihisa Tanaka, Yasushi Iimura, Takumi Mitsuhashi, Hidenori Sugano, Duo Wang, Md. Khademul Islam Molla, "Statistical Features in High-Frequency Bands of Interictal iEEG Work Efficiently in Identifying the Seizure Onset Zone in Patients with Focal Epilepsy." **Entropy**, Vol. 22, pp. 1415, December, 2020 **(IF=2.494)**.

2. **Most. Sheuli Akter**, Md. Rabiul Islam, Y. Iimura, H. Sugano, K. Fukumori, D. Wang, T. Tanaka, A. Cichocki, "Multiband entropy-based feature-extraction method for automatic identification of epileptic focus based on high-frequency components in interictal iEEG," **Scientific Reports (Nature)**, Vol. 10, pp. 1–17, April, 2020 **(IF=3.998)**.

Conference Papers

1. **Most. Sheuli Akter**, Md. Rabiul Islam, K. Fukumori, Yasushi Iimura, Hidenori Sugano and Toshihisa Tanaka, "Automatic Detection of Epileptic Focus in Ripple and Fast Ripple Bands of Interictal iEEG based on multi-band analysis," **International Conference on Artificial Intelligence in Information and Communication (ICAIIC)**, February 2020, Fukuoka, Japan.

2. **Most. Sheuli Akter**, Md. Rabiul Islam, Toshihisa Tanaka, Kosuke Fukumori, Yasushi Iimura, and Hidenori Sugano, "Automatic identification of epileptic focus on high-frequency components in interictal iEEG," **8th International Congress on Advanced Applied Informatics (IIAI AAI)**, July 2019, Toyama, Japan.

3. M. R. Islam, T. Tanaka, M. K. I. Molla, and **Most. Sheuli Akter**, "Classification of Motor Imagery BCI Using Multiband Tangent Space Mapping," **22nd IEEE Int. Conf. on Digital Signal Processing (DSP)**, London, U.K, Aug 2017, pp. 1-5.

Abstract

1. Toshihisa Tanaka, Hidenori Sugano, Madoka Nakajima, Yasushi Iimura, Higo T, Takumi Mitshhashi, **Akter Most Sheuli**, Md. Rabiul Islam, Fukumori K., "Machine learning for epileptic focus detection using multi band entropy-based feature-extraction method in patients with focal cortical dysplasia," **International Epilepsy Congress (IEC)**, Bangkok, June 2019.

Bibliography

[1] A. Ngugi, S. Kariuki, C. Bottomley, I. Kleinschmidt, J. Sander, and C. Newton, "Incidence of epilepsy", *Neurology*, vol. 77, no. 10, pp. 1005–1012, 2011, ISSN: 0028-3878.

[2] Y. Yang, M. Zhou, Y. Niu, C. Li, R. Cao, B. Wang, P. Yan, Y. Ma, and J. Xiang, "Epileptic seizure prediction based on permutation entropy", *Frontiers in Computational Neuroscience*, vol. 12, p. 55, 2018.

[3] R. S. Fisher, W. V. E. Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel Jr, "Epileptic seizures and epilepsy: Definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe)", *Epilepsia*, vol. 46, no. 4, pp. 470–472, 2005.

[4] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, and J. Xiang, "Epileptic seizure detection based on eeg signals and cnn", *Frontiers in neuroinformatics*, vol. 12, p. 95, 2018.

[5] A. Y. Mutlu, "Detection of epileptic dysfunctions in eeg signals using hilbert vibration decomposition", *Biomedical Signal Processing and Control*, vol. 40, pp. 33–40, 2018.

[6] F. Rosenow and H. Lüders, "Presurgical evaluation of epilepsy", *Brain*, vol. 124, no. 9, pp. 1683–1700, 2001.

[7] K. Klein and F Rosenow, "The epileptogenic zone: General principles", in *Textbook of Epilepsy Surgery*, CRC Press, 2008, pp. 939–941.

[8] A. Pitkänen and T. P. Sutula, "Is epilepsy a progressive disorder? prospects for new therapeutic approaches in temporal-lobe epilepsy", *The Lancet Neurology*, vol. 1, no. 3, pp. 173–181, 2002.

[9] G. Avanzini and S. Franceschetti, "Cellular biology of epileptogenesis", *The Lancet Neurology*, vol. 2, no. 1, pp. 33–42, 2003.

[10] Y. Bozzi, S. Casarosa, and M. Caleo, "Epilepsy as a neurodevelopmental disorder", *Frontiers in psychiatry*, vol. 3, p. 19, 2012.

[11] J. I. Sirven, "Epilepsy: A spectrum disorder", *Cold Spring Harbor perspectives in medicine*, vol. 5, no. 9, a022848, 2015.

[12] E. Sherman, "Maximizing quality of life in people living with epilepsy.", *Canadian Journal of Neurological Sciences*, vol. 36, 2009.

[13] B. K. Steiger and H. Jokeit, "Why epilepsy challenges social life", *Seizure*, vol. 44, pp. 194–198, 2017.

[14] A. Siddiqui, R. Kerb, M. E. Weale, U. Brinkmann, A. Smith, D. B. Goldstein, N. W. Wood, and S. M. Sisodiya, "Association of multidrug resistance in epilepsy with a polymorphism in the drug-transporter gene abcb1", *New England Journal of Medicine*, vol. 348, no. 15, pp. 1442–1448, 2003.

[15] W. Löscher and H. Potschka, "Drug resistance in brain diseases and the role of drug efflux transporters", *Nature Reviews Neuroscience*, vol. 6, no. 8, pp. 591–602, 2005.

[16] J. A. French, "Refractory epilepsy: Clinical overview", *Epilepsia*, vol. 48, pp. 3–7, 2007.

[17] S. A. Russ, K. Larson, and N. Halfon, "A national profile of childhood epilepsy and seizure disorder", *Pediatrics*, vol. 129, no. 2, pp. 256–264, 2012.

[18] C. S. Camfield and P. R. Camfield, "Long-term social outcomes for children with epilepsy", *Epilepsia*, vol. 48, pp. 3–5, 2007.

[19] L. S. Vidyaratne and K. M. Iftekharuddin, "Real-time epileptic seizure detection using eeg", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 2146–2156, 2017.

[20] K. E. Misulis, *Atlas of EEG, seizure semiology, and management*. Oxford University Press, 2013.

[21] R. B. Pachori and S. Patidar, "Epileptic seizure classification in eeg signals using second-order difference plot of intrinsic mode functions", *Computer methods and programs in biomedicine*, vol. 113, no. 2, pp. 494–502, 2014.

[22] L. D. Iasemidis, D.-S. Shiau, W. Chaovalitwongse, J. C. Sackellares, P. M. Pardalos, J. C. Principe, P. R. Carney, A. Prasad, B. Veeramani, and K. Tsakalis, "Adaptive epileptic seizure prediction system", *IEEE transactions on biomedical engineering*, vol. 50, no. 5, pp. 616–627, 2003.

[23] C. Stöllberger and J. Finsterer, "Cardiorespiratory findings in sudden unexplained/unexpected death in epilepsy (sudep)", *Epilepsy research*, vol. 59, no. 1, pp. 51–60, 2004.

[24] P. Kwan and M. J. Brodie, "Early identification of refractory epilepsy", *New England Journal of Medicine*, vol. 342, no. 5, pp. 314–319, 2000.

[25] Z. Chen, M. J. Brodie, D. Liew, and P. Kwan, "Treatment outcomes in patients with newly diagnosed epilepsy treated with established and new antiepileptic drugs: A 30-year longitudinal cohort study", *JAMA neurology*, vol. 75, no. 3, pp. 279–286, 2018.

[26] A. L. Gillespie, "Epilepsy and other chronic convulsive diseases: Their causes, symptoms, and treatment", *Edinburgh medical journal*, vol. 11, no. 3, p. 271, 1902.

[27] J. F. Téllez-Zenteno, R. Dhar, and S. Wiebe, "Long-term seizure outcomes following epilepsy surgery: A systematic review and meta-analysis", *Brain*, vol. 128, no. 5, pp. 1188–1198, 2005.

[28] J. Engel, M. P. McDermott, S. Wiebe, J. T. Langfitt, J. M. Stern, S. Dewar, M. R. Sperling, I. Gardiner, G. Erba, I. Fried, *et al.*, "Early surgical therapy for drug-resistant temporal lobe epilepsy: A randomized trial", *Jama*, vol. 307, no. 9, pp. 922–930, 2012.

[29] G. A. Worrell, L. Parish, S. D. Cranstoun, R. Jonas, G. Baltuch, and B. Litt, "High-frequency oscillations and seizure generation in neocortical epilepsy", *Brain*, vol. 127, no. 7, pp. 1496–1506, 2004.

[30] R. G. Andrzejak, K. Schindler, and C. Rummel, "Nonrandomness, nonlinear dependence, and nonstationarity of electroencephalographic recordings from epilepsy patients", *Physical Review E*, vol. 86, no. 4, p. 046 206, 2012.

[31] S. Li, W. Zhou, Q. Yuan, S. Geng, and D. Cai, "Feature extraction and recognition of ictal eeg using emd and svm", *Computers in Biology and Medicine*, vol. 43, no. 7, pp. 807 –816, 2013, ISSN: 0010-4825.

[32] K. M. Hassan, M. R. Islam, T. Tanaka, and M. K. I. Molla, "Epileptic seizure detection from eeg signals using multiband features with feedforward neural network", in *2019 International Conference on Cyberworlds (CW)*, 2019, pp. 231–238.

[33] Y. Khan and J. Gotman, "Wavelet based automatic seizure detection in intracerebral electroencephalogram", *Clinical Neurophysiology*, vol. 114, no. 5, pp. 898 –908, 2003, ISSN: 1388-2457.

[34] N. Nicolaou and J. Georgiou, "Detection of epileptic electroencephalogram based on permutation entropy and support vector machines", *Expert Systems with Applications*, vol. 39, no. 1, pp. 202–209, 2012, ISSN: 0957-4174.

[35] L. Guo, D. Rivero, and A. Pazos, "Pileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks", *Journal of Neuroscience Methods*, vol. 193, no. 1, pp. 156 – 163, 2010, ISSN: 0165-0270.

[36] A. Hamad, E. H. Houssein, A. E. Hassanien, and A. A. Fahmy, "A hybrid eeg signals classification approach based on grey wolf optimizer enhanced svms for epileptic detection", in *International Conference on Advanced Intelligent Systems and Informatics*, Springer, 2017, pp. 108–117.

[37] V Sharathappriyaa, S Gautham, and R Lavanya, "Auto-encoder based automated epilepsy diagnosis", in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, 2018, pp. 976–982.

[38] I. Ullah, M. Hussain, H. Aboalsamh, *et al.*, "An automated system for epilepsy detection using eeg brain signals based on deep learning approach", *Expert Systems with Applications*, vol. 107, pp. 61–71, 2018.

[39] R. Sharma, R. B. Pachori, and U. R. Acharya, "Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals", *Entropy*, vol. 2, no. 17, pp. 669–691, 2015.

[40] T. Itakura and T. Tanaka, "Epileptic focus localization based on bivariate empirical mode decomposition and entropy", in *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, IEEE, 2017, pp. 1426–1429.

[41] N Arunkumar, K. Ramkumar, V Venkatraman, E. Abdulhay, S. L. Fernandes, S. Kadry, and S. Segal, "Classification of focal and non focal eeg using entropies", *Pattern Recognition Letters*, vol. 94, pp. 112–117, 2017.

[42] Y. Varatharajah, B. Berry, J. Cimbalnik, V. Kremen, J. V. Gompel, M. Stead, B. Brinkmann, R. Iyer, and G. Worrell, "Integrating artificial intelligence with real-time intracranial EEG monitoring to automate interictal identification of seizure onset zones in focal epilepsy", *Journal of Neural Engineering*, vol. 15, no. 4, p. 046 035, Jun. 2018.

[43] A. Bhattacharyya, M. Sharma, R. B. Pachori, P. Sircar, and R. Acharya, "A novel approach for automated detection of focal EEG signals using empirical wavelet transform", *Neural Computing and applications*, vol. 29, pp. 47–57, 2018.

[44] R. Sharma and R. B. Pachori, "Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions", *Expert Systems with Applications*, vol. 42, no. 3, pp. 1106–1117, 2015, ISSN: 0957-4174.

[45] P. Allen, D. Fish, and S. Smith, "Very high-frequency rhythmic activity during seeg suppression in frontal lobe epilepsy", *Electroencephalography and clinical neurophysiology*, vol. 82, no. 2, pp. 155–159, 1992.

[46] E. Urrestarazu, J. D. Jirsch, P. LeVan, and J. Hall, "High-frequency intracerebral eeg activity (100–500 hz) following interictal spikes", *Epilepsia*, vol. 47, no. 9, pp. 1465–1476, 2006.

[47] M. Zijlmans, P. Jiruska, R. Zelmann, F. S. Leijten, J. G. Jefferys, and J. Gotman, "High-frequency oscillations as a new biomarker in epilepsy", *Annals of neurology*, vol. 71, no. 2, pp. 169–178, 2012.

[48] J. Jacobs, P. LeVan, R. Chander, J. Hall, F. Dubeau, and J. Gotman, "Interictal high-frequency oscillations (80–500 hz) are an indicator of seizure onset areas independent of spikes in the human epileptic brain", *Epilepsia*, vol. 49, no. 11, pp. 1893–1907, 2008.

[49] M. Dümpelmann, J. Jacobs, K. Kerber, and A. Schulze-Bonhage, "Automatic 80–250 hz "ripple" high frequency oscillation detection in invasive subdural grid and strip recordings in epilepsy by a radial basis function neural network", *Clinical Neurophysiology*, vol. 123, no. 9, pp. 1721–1731, 2012.

[50] K. Kerber, M. Dümpelmann, B. Schelter, P. Le Van, R. Korinthenberg, A. Schulze-Bonhage, and J. Jacobs, "Differentiation of specific ripple patterns helps to identify epileptogenic areas for surgical procedures", *Clinical Neurophysiology*, vol. 125, no. 7, pp. 1339–1345, 2014.

[51] R. Zuo, J. Wei, X. Li, C. Li, C. Zhao, Z. Ren, Y. Liang, X. Geng, C. Jiang, X. Yang, and X. Zhang, "Automated detection of high-frequency oscillations in epilepsy based on a convolutional neural network", *Frontiers in Computational Neuroscience*, vol. 13, p. 6, 2019, ISSN: 1662-5188.

[52] M. Navarrete, C. Alvarado-Rojas, M. Le Van Quyen, and M. Valderrama, "Ripplelab: A comprehensive application for the detection, analysis and classification of high frequency oscillations in electroencephalographic signals", *PLOS ONE*, vol. 11, no. 6, pp. 1–27, Jun. 2016.

[53] M. S. Akter, M. R. Islam, Y Iimura, H. Sugano, K. Fukumori, D. Wang, T. Tanaka, and A. Cichocki, "Multiband entropy-based feature-extraction method for automatic identification of epileptic focus based on high-frequency components in interictal ieeg", *Sci Rep (in presss)*,

[54] J. M. Yentes, N. Hunt, K. K. Schmid, P. J. Kaipust, D. McGrath, and N Stergiou, "The appropriate use of approximate entropy and sample entropy with short data sets", *Ann Biomed Eng*, vol. 41, pp. 349–365, 2013.

[55] C. C. Mayer, M. Bachler, M. Hörtenhuber, C. Stocker, A. Holzinger, and S. Wassertheurer, "Selection of entropy-measure parameters for

knowledge discovery in heart rate variability data", *BMC Bioinformatics*, vol. 15, no. S6, S2, 2014.

[56] M. S. Akter, M. R. Islam, K. Fukumori, Y. Iimura, H. Sugano, and T. Tanaka, "Automatic detection of epileptic focus in ripple and fast ripple bands of interictal ieeg based on multi-band analysis", in *2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2020, pp. 490–493.

[57] J Gotman, "Automatic recognition of epileptic seizures in the eeg", *Electroencephalography and Clinical Neurophysiology*, vol. 54, no. 5, pp. 530 –540, 1982, ISSN: 0013-4694.

[58] F. Rosenow and H. Lüders, "Presurgical evaluation of epilepsy", *Brain*, vol. 124, no. 9, pp. 1683–1700, 2001.

[59] H Stefan, C Hummel, G Scheler, A Genow, K Druschky, C Tilz, M Kaltenhäuser, R Hopfengärtner, M Buchfelder, and J Romstöck, "Magnetic brain source imaging of focal epileptic activity: A synopsis of 455 cases", *Brain*, vol. 126, no. 11, pp. 2396–2405, 2003.

[60] L. Jäger, K. J. Werhahn, A. Hoffmann, S. Berthold, V. Scholz, J. Weber, S. Noachtar, and M. Reiser, "Focal epileptiform activity in the brain: Detection with spike-related functional mr imaging—preliminary results", *Radiology*, vol. 223, no. 3, pp. 860–869, 2002.

[61] H. O. Lüders, I. Najm, D. Nair, P. Widdess-Walsh, and W. Bingman, "The epileptogenic zone: General principles", *Epileptic disorders*, vol. 8, no. 2, pp. 1–9, 2006.

[62] J. Talairach and J. Bancaud, "Lesion," irritative" zone and epileptogenic focus", *Stereotactic and Functional Neurosurgery*, vol. 27, no. 1-3, pp. 91–94, 1966.

[63] H. Wieser and D Zumsteg, "Indications for surgical management of epilepsy", in *Textbook of Stereotactic and Functional Neurosurgery*.

[64] M. J. Brodie, S. M. Zuberi, I. E. Scheffer, and R. S. Fisher, "The 2017 ilae classification of seizure types and the epilepsies: What do people with epilepsy and their caregivers need to know?", *Epileptic Disorders*, vol. 20, no. 2, pp. 77–87, 2018.

[65] R. S. Fisher, J. H. Cross, J. A. French, N. Higurashi, E. Hirsch, F. E. Jansen, L. Lagae, S. L. Moshé, J. Peltola, E. Roulet Perez, *et al.*, "Operational classification of seizure types by the international league against epilepsy: Position paper of the ilae commission for classification and terminology", *Epilepsia*, vol. 58, no. 4, pp. 522–530, 2017.

[66] R. S. Fisher, J. H. Cross, C. D'souza, J. A. French, S. R. Haut, N. Higurashi, E. Hirsch, F. E. Jansen, L. Lagae, S. L. Moshé, *et al.*, "Instruction manual for the ilae 2017 operational classification of seizure types", *Epilepsia*, vol. 58, no. 4, pp. 531–542, 2017.

[67] I. E. Scheffer, S. Berkovic, G. Capovilla, M. B. Connolly, J. French, L. Guilhoto, E. Hirsch, S. Jain, G. W. Mathern, S. L. Moshé, *et al.*, "Ilae classification of the epilepsies: Position paper of the ilae commission for classification and terminology", *Epilepsia*, vol. 58, no. 4, pp. 512–521, 2017.

[68] R. S. Chang, C. Y. W. Leung, C. C. A. Ho, and A. Yung, "Classifications of seizures and epilepsies, where are we?–a brief historical review and update", *Journal of the Formosan Medical Association*, vol. 116, no. 10, pp. 736–741, 2017.

[69] M. Hallett, "Myoclonus: Relation to epilepsy", *Epilepsia*, vol. 26, S67–S77, 1985.

[70] E. Niedermeyer and F. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[71] J. Jirsch, E Urrestarazu, P LeVan, A Olivier, F Dubeau, and J Gotman, "High-frequency oscillations during human focal seizures", *Brain*, vol. 129, no. 6, pp. 1593–1608, 2006.

[72] J. X. Tao, A. Ray, S. Hawes-Ebersole, and J. S. Ebersole, "Intracranial eeg substrates of scalp eeg interictal spikes", *Epilepsia*, vol. 46, no. 5, pp. 669–676, 2005.

[73] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design", *IEEE transactions on signal processing*, vol. 40, no. ARTICLE, pp. 2207–2232, 1992.

[74] P. Regalia, *Adaptive IIR Filtering in Signal Processing and Control*. CRC Press, 1995.

[75] H. Ocak, "Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy", *Expert Systems with Applications*, vol. 36, no. 2, pp. 2027–2036, 2009.

[76] P. Deshmukh, R. Ingle, V. Kehri, and R. Awale, "Epileptic seizure detection using discrete wavelet transform based support vector machine", in *2017 International Conference on Communication and Signal Processing (ICCSP)*, IEEE, 2017, pp. 1933–1937.

[77] O. Salem, A. Naseem, and A. Mehaoua, "Epileptic seizure detection from eeg signal using discrete wavelet transform and ant colony classifier", in *2014 IEEE International Conference on Communications (ICC)*, IEEE, 2014, pp. 3529–3534.

[78] R. X. Gao and R. Yan, "From fourier transform to wavelet transform: A historical perspective", in *Wavelets: Theory and Applications for Manufacturing*. Boston, MA: Springer US, 2011, pp. 17–32, ISBN: 978-1-4419-1545-0.

[79] R. Sharma and R. B. Pachori, "Classification of epileptic seizures in eeg signals based on phase space representation of intrinsic mode functions", *Expert Systems with Applications*, vol. 42, no. 3, pp. 1106–1117, 2015.

[80] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis", *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[81] G. Rilling, P. Flandrin, P. Gonçalves, and J. M. Lilly, "Bivariate empirical mode decomposition", *IEEE signal processing letters*, vol. 14, no. 12, pp. 936–939, 2007.

[82] Z. Jun, L. Ming, L. Yuming, and Y. Xiaoqin, "Fatigue fracture image segmentation based on bemd", *Failure Analysis and Prevention*, vol. 6, pp. 70–73, 2011.

[83] X. Xu, H. Li, and A. N. Wang, "The application of bemd to multispectral image fusion", in *2007 International Conference on Wavelet Analysis and Pattern Recognition*, IEEE, vol. 1, 2007, pp. 448–452.

[84] W. Huang and Y. Sun, "A new image watermarking algorithm using bemd method", in *2007 International Conference on Communications, Circuits and Systems*, IEEE, 2007, pp. 588–592.

[85] W. Yang, R. Court, P. J. Tavner, and C. J. Crabtree, "Bivariate empirical mode decomposition and its contribution to wind turbine condition monitoring", *Journal of Sound and Vibration*, vol. 330, no. 15, pp. 3766–3782, 2011.

[86] H. Chen, J. Shen, W. Chen, C. Wu, C. Huang, Y. Yi, and J. Qian, "The bivariate empirical mode decomposition and its contribution to grinding chatter detection", *Applied Sciences*, vol. 7, no. 2, p. 145, 2017.

[87] K. Fukunaga, *Introduction to statistical pattern recognition*. Elsevier, 2013.

[88] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[89] T Hastie, R Tibshirani, and J. Friedman, "The elements of statistical learning: Data mining, inference, and prediction. springer, new york, ny.", 2001.

[90] Q. Gu, Z. Li, and J. Han, "Linear discriminant dimensionality reduction", in *Joint European conference on machine learning and knowledge discovery in databases*, Springer, 2011, pp. 549–564.

[91] I. Blümcke, M. Thom, E. Aronica, D. D. Armstrong, H. V. Vinters, A. Palmini, T. S. Jacques, G. Avanzini, A. J. Barkovich, G. Battaglia, *et al.*, "The clinicopathologic spectrum of focal cortical dysplasias: A consensus classification proposed by an ad hoc task force of the ilae diagnostic methods commission 1", *Epilepsia*, vol. 52, no. 1, pp. 158–174, 2011.

[92] B. Boashash, *Time-frequency signal analysis and processing: a comprehensive reference*. Academic Press, 2015.

[93] K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b", *Frontiers in neuroscience*, vol. 6, p. 39, 2012.

[94] M. R. Islam, M. K. I. Molla, M. Nakanishi, and T. Tanaka, "Unsupervised frequency-recognition method of ssveps using a filter bank implementation of binary subband cca", *Journal of neural engineering*, vol. 14, no. 2, p. 026 007, 2017.

[95] M. R. Islam, T. Tanaka, and M. K. I. Molla, "Multiband tangent space mapping and feature selection for classification of eeg during motor imagery", *Journal of neural engineering*, vol. 15, no. 4, p. 046 021, 2018.

[96] H. Higashi and T. Tanaka, "Simultaneous design of fir filter banks and spatial patterns for eeg signal classification", *IEEE transactions on biomedical engineering*, vol. 60, no. 4, pp. 1100–1110, 2012.

[97] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, "High-speed spelling with a noninvasive brain–computer interface", *Proceedings of the national academy of sciences*, vol. 112, no. 44, E6058–E6067, 2015.

[98] S. Pincus, "Approximate entropy (apen) as a complexity measure", *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 5, no. 1, pp. 110–117, 1995.

[99] D. Abásolo, R. Hornero, P. Espino, J. Poza, C. I. Sánchez, and R. de la Rosa, "Analysis of regularity in the eeg background activity of alzheimer's disease patients with approximate entropy", *Clinical Neurophysiology*, vol. 116, no. 8, pp. 1826–1834, 2005.

[100] U. R. Acharya, O. Faust, V. Sree, G Swapna, R. J. Martis, N. A. Kadri, and J. S. Suri, "Linear and nonlinear analysis of normal and cad-affected heart rate signals", *Computer methods and programs in biomedicine*, vol. 113, no. 1, pp. 55–68, 2014.

[101] 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, vol. 278, no. 6, H2039–H2049, 2000.

[102] D. Cui, J. Wang, Z. Bian, Q. Li, L. Wang, and X. Li, "Analysis of entropies based on empirical mode decomposition in amnesic mild cognitive impairment of diabetes mellitus", *Journal of Innovative Optical Health Sciences*, vol. 8, no. 05, p. 1 550 010, 2015.

[103] M. O. Sokunbi, V. B. Gradin, G. D. Waiter, G. G. Cameron, T. S. Ahearn, A. D. Murray, D. J. Steele, and R. T. Staff, "Nonlinear complexity analysis of brain fmri signals in schizophrenia", *Plos one*, vol. 9, no. 5, e95146, 2014.

[104] C. Bandt and B. Pompe, "Permutation entropy: A natural complexity measure for time series", *Physical review letters*, vol. 88, no. 17, p. 174 102, 2002.

[105] A. L. Vanluchene, H. Vereecke, O. Thas, E. P. Mortier, S. L. Shafer, and M. M. Struys, "Spectral entropy as an electroencephalographic measure of anesthetic drug effecta comparison with bispectral index and processed midlatency auditory evoked response", *Anesthesiology: The Journal of the American Society of Anesthesiologists*, vol. 101, no. 1, pp. 34–42, 2004.

[106] S. Blanco, A. Garay, and D. Coulombie, "Comparison of frequency bands using spectral entropy for epileptic seizure prediction", *ISRN neurology*, vol. 2013, 2013.

[107] A. Mirzaei, A. Ayatollahi, P. Gifani, and L. Salehi, "Eeg analysis based on wavelet-spectral entropy for epileptic seizures detection", in *2010 3rd International Conference on Biomedical Engineering and Informatics*, IEEE, vol. 2, 2010, pp. 878–882.

[108] C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra", *IEEE Signal processing magazine*, vol. 10, no. 3, pp. 10–37, 1993.

[109] C. Tsallis, "Computational applications of nonextensive statistical mechanics", *Journal of Computational and Applied Mathematics*, vol. 227, no. 1, pp. 51–58, 2009.

[110] C. E. Shannon, "A mathematical theory of communication", *ACM SIGMOBILE mobile computing and communications review*, vol. 5, no. 1, pp. 3–55, 2001.

[111] Y. Chen, P Lin, J. He, Y He, and X. Li, "Combination of the manifold dimensionality reduction methods with least squares support vector machines for classifying the species of sorghum seeds", *Scientific reports*, vol. 6, p. 19 917, 2016.

[112] J. P. Cunningham and M. Y. Byron, "Dimensionality reduction for large-scale neural recordings", *Nature neuroscience*, vol. 17, no. 11, pp. 1500–1509, 2014.

[113] T. Uehara, M. Sartori, T. Tanaka, and S. Fiori, "Robust averaging of covariances for eeg recordings classification in motor imagery brain-computer interfaces", *Neural computation*, vol. 29, no. 6, pp. 1631–1666, 2017.

[114] L. Clemmensen, T. Hastie, D. Witten, and B. Ersbøll, "Sparse discriminant analysis", *Technometrics*, vol. 53, no. 4, pp. 406–413, 2011.

[115] Y. Zhang, G. Zhou, J. Jin, Q. Zhao, X. Wang, and A. Cichocki, "Aggregation of sparse linear discriminant analyses for event-related potential classification in brain-computer interface", *International journal of neural systems*, vol. 24, no. 01, p. 1 450 003, 2014.

[116] Z. Lei, S. Liao, and S. Z. Li, "Efficient feature selection for linear discriminant analysis and its application to face recognition", in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, IEEE, 2012, pp. 1136–1139.

[117] S. Z. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images", *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, no. 4, pp. 627–639, 2007.

[118] K. Sjöstrand, L. H. Clemmensen, R. Larsen, G. Einarsson, and B. K. Ersbøll, "Spasm: A matlab toolbox for sparse statistical modeling", *Journal of Statistical Software*, vol. 84, no. 10, 2018.

[119] H. He, Y. Bai, E. A. Garcia, and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning", in *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*, IEEE, 2008, pp. 1322–1328.

[120] H. He and E. A. Garcia, "Learning from imbalanced data", *IEEE Transactions on knowledge and data engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[121] S. Wang, L. L. Minku, and X. Yao, "A systematic study of online class imbalance learning with concept drift", *IEEE transactions on neural networks and learning systems*, vol. 29, no. 10, pp. 4802–4821, 2018.

[122] B Santoso, H Wijayanto, K. Notodiputro, and B Sartono, "Synthetic over sampling methods for handling class imbalanced problems: A review", in *IOP conference series: earth and environmental science*, vol. 58, 2017, p. 012 031.

[123] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique", *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[124] T. Fawcett, "Roc graphs: Notes and practical considerations for researchers", *Machine learning*, vol. 31, no. 1, pp. 1–38, 2004.

[125] H. Guo and H. L. Viktor, "Learning from imbalanced data sets with boosting and data generation: The databoost-im approach", *ACM Sigkdd Explorations Newsletter*, vol. 6, no. 1, pp. 30–39, 2004.

[126] M. Kubat, R. C. Holte, and S. Matwin, "Machine learning for the detection of oil spills in satellite radar images", *Machine learning*, vol. 30, no. 2-3, pp. 195–215, 1998.

[127] M. Kubat, S. Matwin, *et al.*, "Addressing the curse of imbalanced training sets: One-sided selection", in *Icml*, Citeseer, vol. 97, 1997, pp. 179–186.

[128] T. Fawcett, "An introduction to roc analysis", *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861 –874, 2006, ROC Analysis in Pattern Recognition, ISSN: 0167-8655.

[129] A. Medvedev, G. Agoureeva, and A. Murro, "A long short-term memory neural network for the detection of epileptiform spikes and high frequency oscillations", *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.

[130] K. Edakawa, T. Yanagisawa, H. Kishima, R. Fukuma, S. Oshino, H. M. Khoo, M. Kobayashi, M. Tanaka, and T. Yoshimine, "Detection of epileptic seizures using phase–amplitude coupling in intracranial electroencephalography", *Scientific reports*, vol. 6, p. 25 422, 2016.

[131] N. E. Cámpora, C. J. Mininni, S. Kochen, and S. E. Lew, "Seizure localization using pre ictal phase-amplitude coupling in intracranial electroencephalography", *Scientific Reports*, vol. 9, no. 1, pp. 1–8, 2019.

[132] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on

recording region and brain state", *Physical Review E*, vol. 64, no. 6, p. 061 907, 2001.

[133] V. Bajaj and R. B. Pachori, "Epileptic seizure detection based on the instantaneous area of analytic intrinsic mode functions of eeg signals", *Biomedical Engineering Letters*, vol. 3, no. 1, pp. 17–21, 2013.

[134] K. Sato, N. Arai, A. Omori-Mitsue, A. Hida, A. Kimura, and S. Takeuchi, "The prehospital predictors of tracheal intubation for in patients who experience convulsive seizures in the emergency department", *Internal Medicine*, vol. 56, no. 16, pp. 2113–2118, 2017.

[135] P. Boonluksiri, A. Visuthibhan, and K. Katanyuwong, "Clinical prediction rule of drug resistant epilepsy in children", *Journal of epilepsy research*, vol. 5, no. 2, p. 84, 2015.

[136] C. F. ANDERSON, K. MOXNESS, J. MEISTER, and M. F. BURRITT, "The sensitivity and specificity of nutrition-related variables in relationship to the duration of hospital stay and the rate of complications", in *Mayo Clinic Proceedings*, Elsevier, vol. 59, 1984, pp. 477–483.

[137] M. Mursalin, Y. Zhang, Y. Chen, and N. V. Chawla, "Automated epileptic seizure detection using improved correlation-based feature selection with random forest classifier", *Neurocomputing*, vol. 241, pp. 204–214, 2017.

[138] D. Wang, D. Miao, and C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical eeg classification for epileptic detection", *Expert Systems with Applications*, vol. 38, no. 11, pp. 14 314–14 320, 2011.

[139] R. Sharma, R. B. Pachori, and U. R. Acharya, "An integrated index for the identification of focal electroencephalogram signals using discrete wavelet transform and entropy measures", *Entropy*, vol. 17, no. 8, pp. 5218–5240, 2015.

[140] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection", in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 975–982.

[141] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using eeg signals", *Computers in biology and medicine*, vol. 100, pp. 270–278, 2018.

[142] J. A. Blanco, M. Stead, A. Krieger, W. Stacey, D. Maus, E. Marsh, J. Viventi, K. H. Lee, R. Marsh, B. Litt, *et al.*, "Data mining neocortical high-frequency oscillations in epilepsy and controls", *Brain*, vol. 134, no. 10, pp. 2948–2959, 2011.

[143] M. Zijlmans, J. Jacobs, Y. U. Kahn, R. Zelmann, F. Dubeau, and J. Gotman, "Ictal and interictal high frequency oscillations in patients with focal epilepsy", *Clinical neurophysiology*, vol. 122, no. 4, pp. 664–671, 2011.

[144] J. Jacobs, R. Zelmann, J. Jirsch, R. Chander, C.-É. C. F. Dubeau, and J. Gotman, "High frequency oscillations (80–500 hz) in the preictal period in patients with focal seizures", *Epilepsia*, vol. 50, no. 7, pp. 1780–1792, 2009.

[145] R. J. Staba, C. L. Wilson, A. Bragin, I. Fried, and J. Engel Jr, "Quantitative analysis of high-frequency oscillations (80–500 hz) recorded in human epileptic hippocampus and entorhinal cortex", *Journal of neurophysiology*, vol. 88, no. 4, pp. 1743–1752, 2002.

[146] A. B. Gardner, G. A. Worrell, E. Marsh, D. Dlugos, and B. Litt, "Human and automated detection of high-frequency oscillations in clinical intracranial eeg recordings", *Clinical neurophysiology*, vol. 118, no. 5, pp. 1134–1143, 2007.

[147] C. Jiang, X. Li, J. Yan, T. Yu, X. Wang, Z. Ren, D. Li, C. Liu, W. Du, X. Zhou, *et al.*, "Determining the quantitative threshold of high-frequency oscillation distribution to delineate the epileptogenic zone by automated detection", *Frontiers in neurology*, vol. 9, p. 889, 2018.

[148] S. Liu, C. Gurses, Z. Sha, M. M. Quach, A. Sencer, N. Bebek, D. J. Curry, S. Prabhu, S. Tummala, T. R. Henry, *et al.*, "Stereotyped high-frequency oscillations discriminate seizure onset zones and critical functional cortex in focal epilepsy", *Brain*, vol. 141, no. 3, pp. 713–730, 2018.

[149] N. Jrad, A. Kachenoura, I. Merlet, F. Bartolomei, A. Nica, A. Biraben, and F. Wendling, "Automatic detection and classification of high-frequency oscillations in depth-eeg signals", *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2230–2240, 2017.

[150] D. Lai, X. Zhang, K. Ma, Z. Chen, W. Chen, H. Zhang, H. Yuan, and L. Ding, "Automated detection of high frequency oscillations in intracranial eeg using the combination of short-time energy and convolutional neural networks", *IEEE Access*, vol. 7, pp. 82 501–82 511, 2019.

[151] I Blu, M Thom, E Aronica, *et al.*, "The clinico-pathological spectrum of focal cortical dysplasias: A consensus classification proposed by an ad hoc task force of the ilae diagnostic methods commission", *Epilepsia*, vol. 52, pp. 158–74, 2011.

[152] Y. Liu, W. Zhou, Q. Yuan, and S. Chen, "Automatic seizure detection using wavelet transform and svm in long-term intracranial eeg", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 6, pp. 749–755, 2012.

[153] J. Yoo, L. Yan, D. El-Damak, M. A. B. Altaf, A. H. Shoeb, and A. P. Chandrakasan, "An 8-channel scalable eeg acquisition soc with patient-specific seizure classification and recording processor", *IEEE Journal of Solid-State Circuits*, vol. 48, no. 1, pp. 214–228, 2013.

[154] S. M. S. Alam and M. I. H. Bhuiyan, "Detection of epileptic seizures using chaotic and statistical features in the emd domain", in *2011 Annual IEEE India Conference*, 2011, pp. 1–4.

[155] L. Wang, W. Xue, Y. Li, M. Luo, J. Huang, W. Cui, and C. Huang, "Automatic epileptic seizure detection in eeg signals using multi-domain feature extraction and nonlinear analysis", *Entropy*, vol. 6, no. 19, p. 222, 2017.

[156] R. J. Staba, C. L. Wilson, A. Bragin, I. Fried, and J. Engel, "Quantitative analysis of high-frequency oscillations (80–500 hz) recorded in human epileptic hippocampus and entorhinal cortex", *Journal of Neurophysiology*, vol. 88, no. 4, pp. 1743–1752, 2002.

[157] S. Chaibi, Z. Sakka, T. Lajnef, M. Samet, and A. Kachouri, "Automated detection and classification of high frequency oscillations (hfos) in human intracereberal eeg", *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 927 –934, 2013, ISSN: 1746-8094.

[158] K. S. Kim, H. H. Choi, C. S. Moon, and C. W. Mun, "Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions", *Current Applied Physics*, vol. 11, no. 3, pp. 740 –745, 2011, ISSN: 1567-1739.

[159] J. Too, A. R. Abdullah, and N. M. Saad, "Classification of hand movements based on discrete wavelet transform and enhanced feature extraction", *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 6, 2019.

[160] T. Das, A. Ghosh, S. Guha, and P. Basak, "Classification of eeg signals for prediction of seizure using multi-feature extraction", in *2017 1st International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech)*, 2017, pp. 1–4.

[161] M. K. Hasan, M. A. Ahamed, M. Ahmad, and M. A. Rashid, "Prediction of epileptic seizure by analysing time series eeg signal using k-nn classifier", *applied bionics and biomechanics*, vol. 2017, p. 12, 2017.

[162] W.-T. Shi, Z.-J. Lyu, S.-T. Tang, T.-L. Chia, and C.-Y. Yang, "A bionic hand controlled by hand gesture recognition based on surface emg signals: A preliminary study", *Biocybernetics and Biomedical Engineering*, vol. 38, no. 1, pp. 126–135, 2018.

[163] M. Ariyanto, W. Caesarendra, K. A. Mustaqim, M. Irfan, J. A. Pakpahan, J. D. Setiawan, and A. R. Winoto, "Finger movement pattern recognition method using artificial neural network based on electromyography (emg) sensor", in *2015 International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT)*, 2015, pp. 12–17.

[164] D. Tkach, H. Huang, and T. A. Kuiken, "Study of stability of time-domain features for electromyographic pattern recognition", *Journal of neuroengineering and rehabilitation*, vol. 7, no. 1, p. 21, 2010.

[165] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, "Permutation entropy and its main biomedical and econophysics applications: A review", *Entropy*, vol. 14, no. 8, pp. 1553–1577, 2012.

[166] J. Pohjalainen, O. Räsänen, and S. Kadioglu, "Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits", *Computer Speech and Language*, vol. 29, no. 1, pp. 145 –171, 2015, ISSN: 0885-2308.

[167] C. Yang, R. Duraiswami, and L. S. Davis, "Efficient kernel machines using the improved fast gauss transform", in *Advances in neural information processing systems*, 2005, pp. 1561–1568.

[168] J. Wu, T. Zhou, and T. Li, "Detecting epileptic seizures in eeg signals with complementary ensemble empirical mode decomposition and extreme gradient boosting", *Entropy*, vol. 22, no. 2, p. 140, 2020.

[169] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree", in *Advances in neural information processing systems*, 2017, pp. 3146–3154.

[170] C. Tonini, E. Beghi, A. T. Berg, G. Bogliun, L. Giordano, R. W. Newton, A. Tetto, E. Vitelli, D Vitezic, and S. Wiebe, "Predictors of epilepsy surgery outcome: A meta-analysis", *Epilepsy research*, vol. 62, no. 1, pp. 75–87, 2004.

[171] C. Bergmeir and J. M. Benítez, "On the use of cross-validation for time series predictor evaluation", *Information Sciences*, vol. 191, pp. 192–213, 2012.

[172] L. J. Tashman, "Out-of-sample tests of forecasting accuracy: An analysis and review", *International journal of forecasting*, vol. 16, no. 4, pp. 437–450, 2000.

[173] M. R. Islam, T. Tanaka, and M. K. I. Molla, "Multiband tangent space mapping and feature selection for classification of EEG during motor imagery", *Journal of Neural Engineering*, vol. 15, no. 4, p. 046 021, 2018.

[174] S. J. Pan and Q. Yang, "A survey on transfer learning", *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[175] Y. You, W. Chen, M. Li, T. Zhang, Y. Jiang, and X. Zheng, "Automatic focal and non-focal EEG detection using entropy-based features from

flexible analytic wavelet transform", *Biomedical Signal Processing and Control*, vol. 57, p. 101 761, 2020.

[176] F. Lotte and C. Guan, "Learning from other subjects helps reducing brain-computer interface calibration time", in *2010 IEEE International conference on acoustics, speech and signal processing*, IEEE, 2010, pp. 614–617.