EpiBOX: A Novel Approach to Long-Term Data Acquisition with Automatic Seizure Detection in Epilepsy

Ana Sofia Cacais do Carmo

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisors: Prof. Ana Luísa Nobre Fred
Prof. Hugo Humberto Plácido da Silva

Examination Committee

Chairperson: Prof. Patrícia Margarida Piedade Figueiredo
Supervisor: Prof. Ana Luísa Nobre Fred
Member of the Committee: Prof. Hugo Alexandre Teixeira Duarte Ferreira

January 2021

Dedicated to those who, sometimes, fear they are not enough (including myself).

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Preface

The work presented in this thesis was performed at Institute of Telecommunications, Instituto Superior Técnico (Lisbon, Portugal), during the period February 2020 - January 2021, under the supervision of Prof. Ana Luísa Nobre Fred and Prof. Hugo Plácido da Silva and co-supervised by Dra. Carla Bentes, from Hospital de Santa Maria.

Acknowledgments

I have received unbelievable support during the development of this work and for that I am incredibly thankful.

I would like to start by thanking my parents and my brother, for shaping me to who I am today, for your constant love and attention, and for allowing me to experience every opportunity that led me through this path and culminates with this work. I would also like to thank all other elements of my family: grandparents, cousins, uncles and aunt for supporting me and even trying to understand what this work was about.

I could not go without mentioning some of my closest friends, in particular Joana and Afonso (for sharing with me so many shaping moments, inside and outside of IST), Yasmin (for being with me since forever) and Catarina (for your special energy), as well as the remaining elements of Shishando (for being cool). And of course, a very special appreciation remark for my boyfriend, Lourenço, not only for exceeding at your part of being supportive and caring, but also for being incredibly relevant in all of my journey in IST, without whom it would all have been so much more difficult.

Obviously, this thesis would not be possible without the incredible professionals that continuously invest in research, namely Dr. Carla Bentes and Dra. Ana Rita Peralta from Hospital de Santa Maria and Dr. Franscisca Sá from Hospital Egas Moniz, for handing me their time; as well as the researchers at EmotAI that contributed with one of their devices for this work.

I would also like to thank all fellow researchers at IT, for making me feel welcome, and in particular to Mariana Abreu, for your knowledge, and whose ongoing PhD project on seizure prediction allowed for the framing and evolution of this work. Finally, I would like to thank my thesis supervisor, Ana Fred, for the ideas and the opportunity of working in such an interesting project; and my other supervisor, Hugo Silva, for all the support, feedback and commitment, that exceeded all expectations.

This work has been partially funded by the FCT/MCTES through national funds and when applicable co-funded by EU funds under the project UIDB/50008/2020 ”NICE-HOME”.

I appreciate you all,
Ana Sofia Carmo

Resumo

Epilepsia é uma doença neurológica que afeta cerca de 50 milhões de pessoas em todo o mundo, incluindo 50 mil em Portugal. É uma perturbação do sistema nervoso central, caracterizada por crises recorrentes, que podem ter impactos consideráveis na saúde física e mental dos doentes e das suas famílias.

Investigação é crucial para criar novas e melhores ferramentas de gestão da epilepsia, tanto para os doentes como para os profissionais de saúde. No entanto, dados são fundamentais e nem sempre são acessíveis nas condições ideais. Assim, este trabalho propõe uma alternativa para aquisição de biosinais a partir de *wearables*, com o objetivo final de criar um sistema automático de deteção de crises para a construção de uma base de dados de aquisições de longa duração.

O *EpiBOX* foi concebido como uma ferramenta autónoma de aquisição multimodal de dados, com capacidade para adquirir, reproduzir e armazenar até 12 canais simultaneamente, através de uma interface simples. Adicionalmente, foram desenhados classificadores Support Vector Machines, específicos para oito categorias de crise, utilizando uma configuração limitada de canais (*Fp1-Fp2*). Utilizando a base de dados de TUH EEG Seizure Corpus, foram obtidos resultados excecionais para as crises tónico-clónicas e mioclónicas, com sensibilidades de 98.9% e 98.2%, bem como precisões de 100% e 99.8%, respetivamente.

Esta dissertação apresenta uma base de trabalho importante para outros projetos colaborativos mais envolventes, não só na área de epilepsia como outras, desempenhando tanto um papel de ferramenta auxiliar para investigação, como de primeiros passos para a criação de algumas soluções tecnológicas.

Palavras-chave: epilepsia, deteção de crises, EEG, aquisição de biossinais, BITalino, Raspberry Pi

Abstract

Epilepsy is a neurological disease that affects about 50 million people worldwide, and around 50 thousand in Portugal alone. It is a disorder of the central nervous system, characterized by recurrent seizures that can have a massive impact in the physical and mental health of the people who suffer from it, as well as their loved ones.

Research is an invaluable tool in the improvement of conditions for clinicians and patients to deal with epilepsy. However data is key, and it is not always available, at least not in the desired conditions. Accordingly, this work proposes a practical alternative for the acquisition of biosignals using wearable devices, with the ultimate goal of providing a fully automated seizure detection system, for scalable long-term dataset creation.

EpiBOX was conceived as a practical and standalone multimodal acquisition system with capacity for acquiring, displaying and storing up to 12 different channels, simultaneously, with a simple interaction framework. Additionally, a seizure-specific Support Vector Machines classifier was designed for eight different types of seizure, using a limited-channel configuration (*Fp1-Fp2*). The dataset used was TUH EEG Seizure Corpus, for which phenomenal results were achieved for Tonic-Clonic Seizures and Myoclonic Seizures, with sensitivities of 98.9% and 98.2%, as well as precisions of 100% and 99.8%, respectively.

This dissertation provides important ground work for larger collaborative projects in the field of epilepsy and others, serving both as a complementary tool for research, as well as first steps for some technological solutions.

Keywords: epilepsy, seizure detection, EEG, biosignal acquisition, BITalino, Raspberry Pi

Contents

List of Tables

List of Figures

List of Acronyms

ABSZ Absence Seizures. xviii, 46, 53, 55, 57, 60–64, 83, 84

ACC Accelerometer. 29, 37, 40

Acc Accuracy. 17, 84

ADC Analog-to-Digital Converter. 34

AEDs Anti-Epileptic Drugs. 8, 9

ANS Autonomic Nervous System. 29

ApEn Approximate Entropy. 25, 27

AR Average Reference. xviii, 45, 46, 61, 63

ATSZ Atonic Seizures. xviii, 46

BCKG Background. 46

CHB-MIT CHB-MIT Scalp EEG Database. 2, 28

CNN Convolutional Neural Network. 28

CNSZ Clonic Seizures. xviii, 46

CPSZ Complex Partial Seizures. xviii, 46, 53, 55, 57, 60–64, 83, 84

CWT Continuous Wavelet Transform. 25

db4 Daubechies 4. xviii, 50

DWT Discrete Wavelet Transform. xviii, xxi, xxii, 19, 20, 25, 49–51

ECG Electrocardiography. 2, 29, 30, 34, 38

EDA Electrodermal Activity. 29, 34, 38

EDF European Data Format. xxii, 45, 47

EEG Electroencephalography. xxiii, 2, 3, 9–11, 13, 18, 23–31, 34, 38, 40, 43, 45, 47–49, 59, 62–65, 74

EMG Electromyography. 38

EOG Electrooculography. 38

FD Fractal Dimension. 25, 50

FIR Finite Impulse Response. 19

fNIRS functional Near-Infrared Spectroscopy. 30

FNSZ Focal Non-Specific Seizures. xviii, 46, 53, 57, 60–64, 83, 84

FPR False Positive Rate. xxi, 27, 29–31

FPS Frames per Second. 41

FT Fourier Transform. 25

GA Genetic Algorithm. 26

GNSZ Generalized Seizures. xviii, 46, 53, 57, 60–64, 83, 84

HE Hurst Exponent. 25, 50

HRV Heart Rate Variability. 29

IQR Interquartile Range. 15, 57

IT Instituto de Telecomunicações. 2, 34, 64, 65

KDE Kernel Density Estimate. xxii, 54

KNN K-Nearest Neighbors. 28

LDA Linear Discriminant Analysis. 15, 27, 28

LE Linked-Ears Reference. xviii, 45, 46, 61, 63

LSTM Long Short-Term Memory. 28

ML Machine Learning. xxi, 2, 3, 11–18, 24, 26, 52, 54, 58

MQTT Message Queue Telemetry Transport. xxii, 21, 22, 33, 34, 36, 39, 41, 42, 63, 78

MYSZ Myoclonic Seizures. xiii, xviii, 3, 46, 53, 55, 57, 60–64, 83, 84

NB Naive Bayes. 28

NESZ Non-Epileptic Seizures. 46

P Precision. 17, 84

PPG Photoplethysmography. 29, 37, 40, 43

PRV Pulse Rate Variability. 29

PSD Power Spectral Density. xxii, 25, 47, 48

PZT Respiration. 38

RBF Radial Basis Function. 58, 83

RF Random Forest. 26, 28

RPi Raspberry Pi. xix, xxii, 33, 34, 36, 38, 41, 42, 63, 75, 78, 79

SampEn Sample Entropy. 25, 49, 50

sEMG surface Electromyography. 29

Sn Sensitivity. 17, 84

SNR Signal-to-Noise Ratio. 13, 47

Sp Specificity. 17

SpO_2 Oxygen Saturation. 37, 40

SPSZ Simple Partial Seizures. xviii, 46, 53, 57, 60–64, 83, 84

SUDEP sudden unexpected death in epilepsy. 1, 7

SVM Support Vector Machines. xi, xiii, xxii, 15, 27–29, 58, 59, 83

TCP Temporal Central Parasagittal. 26

TCSZ Tonic-Clonic Seizures. xiii, xviii, xxii, 3, 46, 53–55, 57, 60–64, 83, 84

TD Time Domain. 25

TNSZ Tonic Seizures. xviii, 46, 53, 57, 60–64, 83, 84

TUSZ TUH EEG Seizure Corpus. xi, xiii, 2, 3, 28, 45

UI User Interface. 41, 63, 64, 75

VNS Vagus Nerve Stimulation. 9

WHO World Health Organization. 1, 7

WT Wavelet Transform. 19, 27

Chapter 1

Introduction

1.1 Motivation

According to the World Health Organization (WHO), 50 million people worldwide, and 50 thousand people in Portugal[1], suffer from epilepsy, a noncommunicable neurological disease that can affect people of all ages. Epilepsy is a disorder of the central nervous system, characterized by recurrent seizures, which are sudden and uncontrolled disturbances in the electrical activity of the brain. These seizures result from an "hypersynchronous discharge of a population of cortical neurons" [1] and may result in a large spectrum of symptoms, ranging from temporary confusion to uncontrollable jerking movements of the limbs, loss of consciousness or even sudden unexpected death in epilepsy (SUDEP)[2].

This condition can have tremendous effects on the lives of the people who suffer from it and the people close to them, either due to the health impacts it carries, such as the physical impairments that can result from a seizure and the increased risk of other chronic health conditions, or due to the socio-economic burden it often brings. Several constraints can prevent the patient from living a normal life, including driving limitations (imposed according to the nature of the seizures) that can easily limit access to healthcare facilities and job opportunities if no other means of transportation are available. These limitations only further enhance the patient's dependency on others, which can evidently impact their self-confidence [2]. Moreover, stigma, which can arise from misconceptions and fear around epilepsy, can result in discrimination, rejection and negative self image which greatly contribute to the devastating effects of this condition on the mental health of its patients and their relatives. Accordingly, self-management of epilepsy is key, motivating the need for skills/tools that allow patients to control their seizures, prevent injuries, make healthy lifestyle choices and thus taking "control of their health and cope with their day-to-day challenges" [3].

Moreover, the COVID-19 pandemic mitigation measures introduced additional difficulties for the management of the disease. In particular, there has been a significant decrease in the patient-clinician encounters and inpatient visits for long-term monitoring. Clinical exams have been reduced to an in-

[1] J. Pimentel. "O doente epiléptico pode e deve ter uma vida normal". [Online].
[2] Mayo Clinic. "Seizures". [Online].
[3] Managing Epilepsy Well Network. "Epilepsy self-management is a partnership". [Online].

dispensable minimum, and the interactions between patients and researchers in the field of epilepsy have been greatly conditioned. This created a pressing need for new ways of facilitating wearable and practical patient monitoring, especially based on remote and automated minimally-assisted methods.

1.2 Work Overview

Research in biomedical signals and systems has an invaluable role in the evolution of how patients and clinicians deal with epilepsy (as in any other medical condition), as it can provide novel tools to diagnose, monitor and manage the condition. An essential agent in research is data and, in some cases, it is not easily available. In epilepsy, there are several available open-access online datasets (such as CHB-MIT Scalp EEG Database (CHB-MIT) [3] and TUH EEG Seizure Corpus (TUSZ) [4]), however, each dataset has its own singularities, either regarding seizure type, demographics of the subjects, annotations, etc.

Moreover, there is an increasing pursuit of wearable solutions that can be applied in every-day life monitoring, which uncovers a clear weakness in the online datasets: the available data is collected with inpatient, non-portable, devices that are not subject to the same conditions as wearable devices (e.g. artifacts and noise), which means the acquired biosignals do not replicate the signals acquired with wearables, hampering the study of such montages. The use of non-Electroencephalography (non-EEG) biosignals in the study of epilepsy is rapidly gaining ground as well [5], which unveils another weakness in the available datasets: seldom they provide other modalities, such as Electrocardiography (ECG), but these are inconsistent (even within the same dataset) and ECG is usually the only other modality available, which, again, hinders research in this field. Therefore, it is important to provide researchers the possibility of recording their own data and personalize it to their studies and ideas.

Within the research group of Instituto de Telecomunicações (IT), there are several ongoing projects addressing this issue, particularly in the field of wearable devices applied to epilepsy, in partnership with Hospital de Santa Maria and Hospital Egas Moniz. However, within the course of this collaborations, a major bottleneck was identified: the need for technical personnel to operate the acquisition system, which evidently restrains the acquisition capabilities and possibility for growth. Moreover, the current acquisition system relies on the constant use of a PC within the hospital wards, which is neither appealing nor liable to be sanitized.

Therefore, the purpose of this work is to create an autonomous system for continuous or near-continuous data acquisition using wearable devices (*EpiBOX*) and a seizure detection algorithm. This is a necessary stepping stone to achieve the ultimate goal of providing a fully automated seizure detection system, which will enable scalable long-term dataset creation containing in-hospital and at-home recordings, both longitudinal and cross-sectional [6]. In long-term recordings, seizure detection is a fundamental component, as it enables automated dataset annotation, thus facilitating the task of identifying the regions of interest within the many hours of recording. Considering that ground-truth is the basis for most Machine Learning (ML) algorithms, seizure detection is complementary to prediction. Moreover, *EpiBOX* is not limited to seizure detection and EEG acquisition. It is designed as a versatile acquisi-

tion system, that can record a multimodal biosignal data, and that can be integrated with whichever ML algorithm.

This work effectively establishes the foundations for the proposed system, in which two main components will be described, corresponding to 1) a proposal of an EEG-based automatic seizure detection algorithm, designed for a limited-channel headband and 2) *EpiBOX* - a standalone acquisition system for multimodal biosignal data acquisition, that can be easily handled through a mobile app.

1.3 Contributions

This work delivers a completely operational long-term acquisition tool, designed as an autonomous, small and versatile system to record multimodal biosignal data, that can be integrated with whichever machine learning algorithm. Moreover, within *EpiBOX*, the mobile application responsible for the user interaction with the software, provides a simple and intuitive way of interaction, enabling the monitoring and acquisition of data by non-technical personnel.

EpiBOX initially arises from a need of a more practical setup for the acquisition of long-term data in the context of an epilepsy-related PhD research. However, it progressed to become a complementary tool for research, not only for EEG or epilepsy, but for every field of study of biosignals.

Furthermore, with this work we showed that it is possible to automatically detect Tonic-Clonic Seizures and Myoclonic Seizures using the limited channel configuration of *Fp1-Fp2*, which to the author's knowledge, had not yet been approached in literature.

1.4 Thesis Outline

This document is divided into six chapters, throughout which the main concepts behind the work are addressed, the developed algorithm and mobile application are explained, and the technical and scientific grounding is provided. In the present chapter, the topic of the thesis is presented, as well as its context and applicability perspectives in real-world settings. Chapter 2 states the background knowledge on epilepsy, electroencephalography and some key concepts regarding machine learning and communication protocols. Chapter 3 presents the state-of-the-art, focusing on the studies performed towards minimizing the number of EEG channels needed for seizure detection, as well as some of the work being done in the scope of multimodal analysis of seizures; it also reviews the current wearable solutions for the acquisition of EEG. Chapter 5 explains the process and methods behind the architecture of the detection algorithm, designed for EmotAI's headband, along with its performance on the TUSZ online dataset. Chapter 4 describes the implementation of *EpiBOX*, coupled with some preliminary reflections on its operation. Finally, Chapter 6 outlines the main achievements and drawbacks of this work followed by some suggestions regarding future work.

Chapter 2

Background

2.1 Background on Epilepsy

2.1.1 Mechanisms of Epilepsy

Epilepsy is a chronic condition of the central nervous system, characterized by recurrent seizures, which are sudden and uncontrolled disturbances in the electrical activity of the brain. However, a single seizure does not imply epilepsy, but rather it is diagnosed only in the event of at least two unprovoked seizures. This disease exhibits a significantly high prevalence over the whole population, with annual incidences of around 40-70/100 000 in Europe and North America and twice the incidence in developing countries (likely due to higher rates of traffic injuries, increased risk of endemic conditions and poorer medical care and infrastructures), making it one of the most prevalent neurological disorders. [1]

Within our nervous system, communication is performed by specialized cells, called neurons, that transmit the information in the form of electrical potentials. This transmission is performed through thin fibers called axons, and the information is shared with neighboring neurons through the release of neurotransmitters into the respective synapses (junctions between communicating neurons) [7].

The excitation and firing of a single neuron, and consequent transmission of an action potential through it, is dependent on a threshold mechanism, and directly affects the membrane potential of the neuron. When a neuron is not activated, the (resting) membrane potential is ≈ -70mV, with about 10x higher concentration of sodium (Na^+) outside the membrane and 30x higher concentration of potassium (K^+) inside. Whenever the magnitude of the potential is large enough to exceed the threshold, the permeability of the membrane to Na^+ and K^+ change, allowing for a rapid influx of Na^+, which induces a positive membrane (action) potential.

The subsequent repolarization and hyperpolarization (decrease in membrane potential to rest state and further, respectively) is achieved with the slower increase of K^+ permeability, that allows for the outflow of these ions, restoring the membrane's resting potential. [8]

It is important to note that the ion permeability is controlled through the excitatory and inhibitory neurotransmitters and their receptors (e.g. Glutamate and GABA, respectively). Outside the action

[1] World Health Organization. "Epilepsy". [Online].

potentials, the resting membrane potential is maintained by pumps present in the membrane (such as the sodium-potassium ATPase), that balance the chemical and electrical gradients, by actively transporting the ions across the membrane. Given this, hyperexcitability of a neuron occurs when there is a disturbance in the balance between excitation and inhibition, i.e. if there is too much excitation or too little inhibition, which can happen if one (or more) of the regulatory mechanisms is disrupted (such as blockage of the GABA inhibitory receptors or abnormalities in ATPase pumps) [9].

This process reflects on the membrane potential of the hyperexcited neurons, as illustrated in Figure 2.1: the depolarization is usually higher in voltage and prolonged in time, during which characteristic mini-spikes appear (which might be responsible for the recruitment of adjacent neurons during a seizure); and the hyperpolarization also has greater magnitude and duration, when compared to a "normal" action potential.

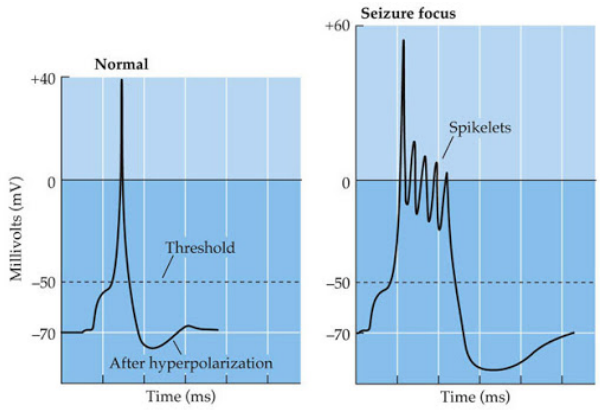


Figure 2.1: Illustration a normal action potential and the counterpart in a neuron within a seizure focus (hyperexcited neuron). Reprinted from [10]

.

However, for seizures to occur, excessive discharge of a single neuron is not enough, as it implies the synchronization of a network of neurons, hence called hypersynchronous discharges. Epileptogenesis conceptualizes the pathogenic process in which a normal neuron network is transformed into an hyperexcitable one, capable of generating spontaneous seizures, and can be characterized by several mechanisms, such as loss of inhibitory neurons (illustrated in Figure 2.2 A), loss of excitatory neurons "driving" inhibitory neurons, pathologies in the ion channels (channelopathies) and excitatory axonal "sprouting" (illustrated in Figure 2.2 B)[2].

It is important to state that not all seizures are epileptic seizures, as they may arise from other phenomena, namely traumatic events such as stroke, brain infections, head injury and birth-related brain injuries. In epileptic seizures in particular, there are three distinctly relevant features: the seizure focus (i.e. the region of the brain where the sudden excessive electrical discharge started), how far it spreads across the brain and the symptoms it causes. In focal (or partial) seizures, the focus is confined to a specific region of the brain (and may spread from there), whereas in generalized seizures the excessive electrical discharge starts virtually simultaneously across the brain. Table 2.1 describes the different seizure types within these two categories, and based on the three key features.

[2]American Epilepsy Society. "Basic mechanisms underlying seizures and epilepsy." [Online].

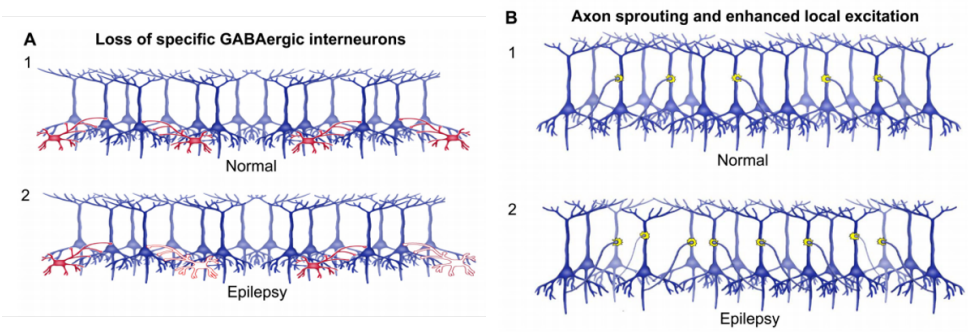


Figure 2.2: (A) Illustration of loss of GABA interneurons during epileptogenesis (B) Illustration of axon sprouting during epileptogenesis. Reprinted from [11]

.

Table 2.1: Types of seizures within focal and generalized seizures, along with a description and common manifestations [12]. The second column corresponds to the prevalence of each seizure type and can be read as "About $x\%$ of people who suffer from epilepsy have this type of seizure". Source: American Epilepsy Society. "Facts and Figures".

Seizure Type	%	Description
Tonic	–	Associated with stiffening of the muscles and impaired consciousness.
Clonic	–	Sustained, rhythmic jerking and potential loss of consciousness, followed by confusion.
Tonic-clonic	25	Involves loss of consciousness and violent muscle contractions.
Atonic	1	Involves the loss of tone of muscles in the body.
Myoclonic	5	Results in brief involuntary twitching or myoclonus.
Absence	5	Involves lapse in attention, may result in impaired memory.
Simple partial	14	Onset in one location of the brain (with the possibility of spreading). Usually brief, characterized by full awareness but may cause sensory responses.
Complex partial	36	Same as simple focal, but characterized by impaired awareness.

By identifying the types of seizures and excluding other causes for them, epilepsy can sometimes be characterized as a syndrome, which is often helpful when deciding on treatment and medication or making prognosis. For this to happen, the disorder must be defined by a group of features, namely causes of seizures, type/types of seizures that occur, onset of epilepsy (age at which is begins), symptoms and signs [3].

2.1.2 Importance of Self-Management and Data-Driven Care

Epilepsy can have several associated comorbidities, namely cognitive impairments, sleep disorders, physical disabilities, migraines and injuries derived from seizure accidents. Besides, the symptoms associated to seizure events can vary from simple lapses of attention and muscle jerks, to prolonged convulsions that can ultimately result in death - sudden unexpected death in epilepsy (SUDEP). According to the WHO, the risk of premature death can be up to three times higher for people suffering from epilepsy, particularly in developing countries. These seizures can also sometimes induce loss of consciousness and control of bowel and bladder function, severely impairing the independence of people in

[3] R.D. Sheth. "EEG in Common Epilepsy Syndromes". [Online].

their day-to-day life. Moreover, stigma, arising from misconceptions and fear around epilepsy, can result in discrimination, rejection and negative self-image, which greatly contribute to the devastating effects of this condition on the mental health of its patients and their relatives.

All things considered, taking an active role in the management of epilepsy, as in many other chronic diseases, is crucial. Understanding their diagnosis, building knowledge around the disease, engaging in treatment decisions and creating support systems play vital roles in empowering patients. By providing patients a sense of control over epilepsy, epilepsy can no longer take control over them.

According to Managing Epilepsy Well Network, self-management in epilepsy comprises three main areas: Treatment Management, Lifestyle Management and Seizure Management. Treatment management implies strictly following the medication schedules, keeping clinic appointments and promoting active communication and discussion with health care professionals. Lifestyle management entails both behavioral changes and environment adjustments, which together work towards minimizing seizure triggers and aggravating factors. Seizure management essentially aims at recognizing and avoiding triggers, and one of the main tools to accomplish this is by keeping seizure diaries. This approach, based on seizure tracking, helps recognize possible triggers or aggravating factors (by annotating what happened around the time of seizures) and patterns (e.g. time of day). Moreover, seizure tracking is not only an invaluable tool for the patient but also for the health care provider, as it gives onsite descriptions of the symptoms (which may help recognize the type of seizure) and thus make more informed decisions on what treatments to prescribe and understand their adequacy and side-effects over treatment administration [4].

Nevertheless, physicians still face challenges when following a patient with epilepsy, because an adequate diagnosis is highly dependent on accurate descriptions of a patient's seizures. However, around 56% of seizures are misreported by patients (increasing to 86% for nocturnal seizures)[5], either due to unawareness of the event (common in absence seizures) or due to loss of consciousness/awareness mid-seizure (in which case, a description of the event becomes more challenging and dependent on the external observation of others). Furthermore, even with accurate descriptions of the seizure events, diagnosing the type of seizure is not always achievable[6], which hinders the process of providing data-driven care.

2.1.3 Treatment

Not every patient diagnosed with epilepsy needs treatment, as long as they are able to accurately identify their seizure triggers and avoid them. However, this is not always possible, in which cases, treatment needs to be implemented and, in most situations, can help decrease the number of seizures or cease them completely. Nowadays, treatments for epilepsy may include [2]:

- Anti-Epileptic Drugs (AEDs), i.e. medication;
- Surgical procedure to remove a small part of the brain that is responsible for the seizures;

[4] Epilepsy Foundation. "Managing your Epilepsy". [Online].
[5] Empatica. "3 challenges physicians face in providing care for their epilepsy patients". [Online].
[6] Epilepsy Foundation. "Diagnosing Epilepsy". [Online].

- Implantation of medical devices to control seizures (e.g. Vagus Nerve Stimulation (VNS) and responsive neurostimulation);
- Special diets, such as the ketogenic diet (low carb, high fat diet), that help control seizures, usually in conjunction with medication.

Medication is one of the most common ways to treat epilepsy, and they provide adequate treatment for about 70% of epilepsy patients, successfully controlling their seizures. However, this implies strict adherence to the medication and its schedule, and may need several iterations to find the suitable combination and dosage of the AEDs, in order to avoid unwanted side-effects or dangerous drug interactions.

In cases of intractable epilepsy (i.e. when several trials of medication were proved inadequate), or when pharmacology therapies are associated with severe side effects, epilepsy surgery is often preferable. This non-pharmacological therapy is based on the removal of a small portion of the brain that is responsible for the onset of the seizures. This procedure is always preceded by some kind of diagnostic tool, such as video-EEG, for "definitive epilepsy syndrome classification and possible localization" [13]. For example, in well-localized mesial temporal lobe epilepsy, epilepsy surgery has been proven to be a superior therapy to AEDs, with significant larger success in seizure elimination (64% of patients vs 8%) [13].

Alternatively, Vagus Nerve Stimulation (VNS) is a type of neuromodulation that works by sending electrical pulses to the brain, through the left Vagus Nerve, acting as a "pacemaker for the brain"[7]. This therapy is usually less effective than surgery for well-localized patients and than AEDs in general, however they provide significantly less side effects, which are usually easier to mitigate than the ones associated to medication, proving to be an adequate solution in some cases of focal epilepsy.

Alternatives to pharmacological treatments are sometimes necessary to ensure an adequate quality of life to patients, as AEDs have the potential to cause adverse side effects which can greatly impact a patient's life (such as drowsiness, fatigue, dizziness, blurry vision and lack of coordination) [13]. Ultimately, choosing which treatments to try, based on the associated risks and success rate, is a decision made collectively by the patient and their physician, and may be a lengthy process to find the best fit for each individual patient. This is where the extreme importance of seizure tracking and biosignal monitoring (e.g. video-EEG, as will be explained in Section 2.2) arises, as these tools provide diagnostic information, treatment effectiveness and reports of unexpected side effects.

2.2 Detection of Epileptic Seizures

2.2.1 Electroencephalograpy (EEG)

Nowadays, video-electroencephalography (video-EEG) remains the golden standard for seizure recording and managing in hospital environments, in conjunction with other (more recent) techniques. EEG is

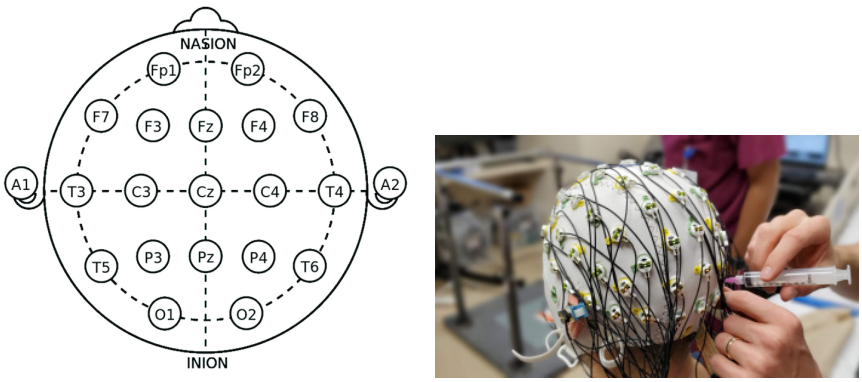
[7] Epilepsy Foundation. "Vagus Nerve Stimulation (VNS)". [Online].

a monitoring technique, with several applications, that records the spontaneous electrical activity of the cerebral cortex, which reflects the summed effects of the numerous excitatory and inhibitory potentials of the cortical neurons [14].

Particularly in epilepsy, during the pre-ictal phase, i.e. before a seizure, the changes in connectivity (inhibitory and excitatory control of neighboring neurons [15]) and activity are reflected as variations in the EEG signal, with characteristic patterns, moments before and during seizures, that are distinct from a non-epileptic signal. Moreover, EEG can also record relevant discrepancies within these abnormal instances, which correspond to the different stages of seizure events: pre-ictal (moments preceding a seizure), ictal (during a seizure) and interictal (between two consecutive seizures, if a patient has multiple within a short period of time) [16].

Hence, EEG is particularly useful in epilepsy as it provides important information to determine seizure focus, allows to distinguish between a focal and generalized seizure disorder, assists in the classification of seizures and, contrary to visual examination of the patient, allows to detect non-convulsive status (seizure with minimal to no motor movements). It gives a powerful insight over the seizure events and may help exclude other conditions that mimic epilepsy.

In clinical setups, EEG is acquired using a head-cap full of electrodes, usually following the international 10-20 system of electrode placement (see Figure 2.3), which has significant coverage across all regions of the scalp. The wave-like patterns seen in EEG result from the potential difference between a pair of electrodes, i.e. each channel does not correspond to the recording of one single electrode, but rather the difference between one of the electrodes and the reference point. This setup allows for great flexibility, as one can play with different montages (arrangements of electrode pairs, called channels) in order to compare the activity in different parts of the brain and thus detect and analyze different epileptiform discharges.



(a) (b)

Figure 2.3: (a) 10-20 international system of electrode placement, with respective electrodes names (Source: Wikipedia, 2010). (b) example of EEG setup in clinical environment, with conductive gel being introduced in the electrodes (Source: Lunatic Laboratories, 2020).

However powerful this technique is, it requires visual inspection of hours-long and multi-channel recordings by experienced professionals, who need to perform careful analysis to adequately identify the ictal and inter-ictal patterns characteristic of epilepsy. To aid this process, video recordings are commonly used in conjunction with EEG (hence why it is called video-EEG) to identify seizure events

by their physical symptoms. This manual annotation process is therefore lengthy, cumbersome and prone to human error, potentially decreasing the accuracy of the annotations, which can then impact the diagnosis (in clinical settings) or the representative samples (in research settings).

In visual analysis of the signal, the professionals are trained to recognize specific patterns (or combinations of patterns) that correspond to the different stages of seizures. These patterns can be spikes (with duration of less than $80ms$), sharp waves (lasting between 80 and $200ms$), polyspikes (series of spikes) or trains of spike-and-wave complexes (usually happen three times per second), but can also be specific rhythms [17]. In normal non-epileptic individuals, these EEG rhythms are usually associated with specific physiological and mental processes and can be more prevalent in particular areas of the brain. Table 2.2 illustrates some of the processes in which each rhythm is more evident and its localization (when applicable).

Table 2.2: EEG frequency bandwidths, corresponding rhythms and association to normal brain activities [18]

Rhythm	**Freq. Range**	**Location**	**Normal Activities**
Delta (δ)	$0.1 \leq f < 4\ Hz$	–	Common in very young children; Some parts of sleep; or Hyperventilation
Theta (θ)	$4 \leq f < 8\ Hz$	–	Common in young children; Drowsiness; or Idling
Alpha (α)	$8 \leq f < 12\ Hz$	Occipital lobes	Adults in relaxed state with eyes closed
Beta (β)	$12 \leq f < 30\ Hz$	Frontal lobes and central area	Awake state
Gamma (γ)	$30 \leq f \leq 70\ Hz$	–	Voluntary motor movement; Learning and memory

In epilepsy, however, these normal patterns are disturbed, not only in ictal segments but also in the EEG background, and each type of epilepsy has its characteristic patterns, which allows physicians to distinguish them.

With the continuous development of knowledge around ML and computer aided diagnostic techniques, researchers quickly attempted to find an automated solution for the cumbersome process of EEG analysis and epileptic seizure detection. The next section introduces some fundamental concepts on Machine Learning, that will be crucial throughout this work.

2.2.2 Fundamentals of Machine Learning

Machine Learning as a Concept

Machine Learning (ML) can be defined as the computational methods that use past evidences from a given process as knowledge to make predictions regarding the categorization of novel (unseen) samples of the same nature [19]. Several tasks can be performed using ML, such as classification (prediction of the class to which the input belongs), regression (prediction of a numerical value) and anomaly detection [20]. Many more tasks exist, but for the purpose of this work, we will focus on the first.

A ML algorithm goes through a set of phases that ultimately result in a model capable of classifying

[8]Generally, it is crucial to split the dataset in train and test first, to avoid placing duplicate samples in both, and consequently causing overfitting and poor generalization (further explained below

new instances of reality. In most setups, at least two disjoint sets of data are necessary[8]: the train set (τ) and the test set (τ'), which allow to train the algorithm and then test its performance on *unseen* data, guaranteeing an unbiased evaluation. In general terms, the ML process follows an architecture similar to the one illustrated in Figure 2.4.

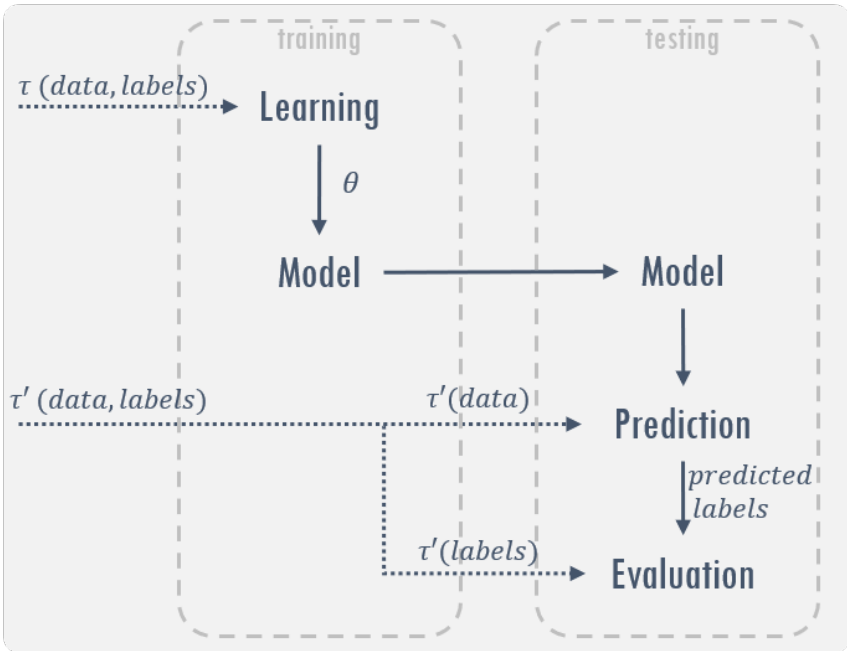


Figure 2.4: Illustration of the general architecture of a ML system. Comprised of a training phase and a testing phase. $\tau(data, labels)$: train set, $\tau'(data, labels)$: test set, θ: model parameters, $predicted\ labels$: predictions for input $\tau'(data)$, which are compared with $\tau'(labels)$ within Evaluation. This diagram does not consider the steps previous to training (e.g. preprocessing and feature extraction), as it is highly dependent on the application.

With some ML models, additional parameters need to be set, called *hyperparameters*, which correspond to extrinsic parameters from the model and can not be estimated from the train data (e.g. number of neighbors, K, in the K-nearest neighbors algorithm). They are used to control the learning process, and can have a massive influence over the performance of the model. These hyperparameters are usually estimated using an heuristic process of grid search (i.e. trial-and-error, going through a list of possible candidate), according to the resulting performance of the model. Therefore, in order to avoid overfitting to the test data, another independent set, known as validation set (τ_v), must be used [21].

In Machine Learning, several scenarios can be employed, such as supervised learning, unsupervised learning and semi-supervised learning. Choosing which scenario to use highly depends on: 1. having labeled training data (i.e. data annotated with the corresponding classes) and 2. the timings on which the train data is available [19].

- *Supervised learning*: training data is available at the beginning of the learning process and all samples are labeled.

- *Unsupervised learning*: training data is available at the beginning of the learning process but no labels are given, hence why learning is highly dependent of the structural organization of the data and the ability to find patterns in it. In such circumstances, performance is usually measured by

intrinsic properties of the resulting clusters, such as cluster separation. This method can also be applied when labels are available if the structure of data encourages it.

- *Semi-supervised learning*: both labeled and unlabeled samples are received (commonly, with a larger number of unlabeled ones), in which case the unlabeled samples are used to reinforce/support the value of the labeled ones.

In practice, many more frameworks exist, but for the purpose of this work, only these will be addressed.

Dataset, Preprocessing and Feature Extraction

One of the key elements of a ML algorithm is the dataset, which comprises the complete set of data (e.g. EEG recordings) that will be used to train and test the algorithm. The original signals usually require preprocessing, which can be seen as a cleaning process to remove irrelevant or distorted data, and may be comprised of filtering, segmentation (i.e. partition of the signal in smaller segments), segment removal, etc.

Preprocessing of biosignals can have a major impact in machine learning applications, as the presence of noise and artifacts can mask relevant information within the signal [7]. Therefore, the removal of these components is key to improve the performance of the system [22]. Many authors have studied and proposed different methods to remove noise and increase the SNR in EEG signals. Most commonly, denoising in biomedical signals is performed (or at least used at some point during the denoising process) through filtering, which allows to eliminate power-line interference ($\approx$ 50 or 60 Hz) and unwanted frequency bands.

From the acquired segments (also called samples or instances), a set features (i.e. attributes of the data) can be extracted, ideally reducing dimensionality while still accurately describing the original dataset. While this step is not suited for some methods (particularly in Deep Learning methods, a subfield of ML), it is key for others, as it allows to extract the most relevant attributes from the data, whilst discarding unwanted/irrelevant ones that would only hinder the performance otherwise. Additional processing can be performed over these features, either at this stage, by identifying redundant features or features with low predictive power; or further along in the process, by evaluating the contribution of each to the performance of the classifier.

Sample Preprocessing

Other processing steps may be necessary in some cases, according to factors relating to the extracted samples. One such example is when we have imbalanced classes (i.e. a disproportionate ratio of samples belonging to each class), which is a common problem in ML. In many tasks, when the purpose is to detect a certain occurrence, this event is commonly less frequent than its counterpart (e.g. a whole-day recording, with a 5 min ictal period, which results in a significantly smaller number of ictal segments, when compared to non-ictal/background segments).

The main issue with this factor is that ML algorithms are designed to maximize a certain metric, such as classification accuracy (refer to the next section for more performance metrics). If the imbalance is very prominent and no measures are taken, maximizing accuracy will correspond to accurately classifying the majority class while completely neglecting the remaining classes, resulting in misleadingly high performances and in a very poor sensitivity to the minority classes[9]. Using different metrics may provide better insight, on the other hand resampling techniques can also provide an adequate solution. When resampling, it is possible to either oversample the minority class (populating the minority class with duplicate or synthetic samples), undersampling the majority class (removing some samples or converging some samples in one), or both.

In some ML methods, other steps may be important such as scaling, which corresponds to applying a transformation individually to each feature in the feature set, to overcome the inherent variety of magnitudes, ranges and units that these features present. Scaling is crucial for distance-based methods, otherwise features with inherently larger ranges would have additional weight in the classification. It may also be advantageous when gradient descent (an optimization algorithm) is used, as it ensures that all step sizes are updated at the same rate and the gradient descent converges more smoothly and quickly towards the minima. Other algorithms, such as tree-based, are virtually insensitive to this processing step.

Two of the most commonly used scaling techniques are normalization (shift and rescale, so that all features range between 0 and 1, (2.1)) and standardization (centering and scaling of each feature to have zero mean and unit standard deviation, (2.2)). Choosing a scaling technique ultimately depends on the task at hand and on the algorithm that will be used and no standard rule exists, however, normalization might be more suited when the distribution of the data does not follow a Gaussian distribution and standardization might be more suited when it does.[10]

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{2.1}$$

$$X' = \frac{X - \mu}{\sigma} \tag{2.2}$$

Outlier removal is another sample processing technique that can be extremely relevant. Outliers represent observations that are too distant from the rest of the sample population and they may correspond to misclassifications of the samples, errors in the acquisition or they might actually just be variance in the data. In the first two cases, the data points result from mistakes, not representing the real environment, and hence do not contribute with any predictive value for the classification. In the last scenario however, they might actually represent another region of the data that we did not expect, and by removing them, an important part of the learning will be lost. Having a large and comprehensive dataset, that represents the largest number of possible events, can help reduce misrepresentations of the feature map and avoid the occurrence of this issue.

Outlier removal is particularly relevant if scaling with a bounding range is performed, to avoid misrepresentation of the data (otherwise all relevant data points will be concentrated in a very restrict range

[9] T. Boyle. "Dealing with imbalanced data". [Online].

[10] A. Bhandari. "Feature Scaling for Machine Learning: Understanding the Difference Between Normalization vs Standardization". [Online].

and the outlier will have more weight in the classification). In order to identify outliers, a number of statistical metrics can be used, such as Z-score and Interquartile Range (IQR) score, which both work on the basis of finding the distribution of the data and setting a threshold, above/below which the data points are considered to be too distant from the mean.

Model Training and Performance Metrics

The training of a ML model can be seen as the process of learning the function that maps the input variables ($\tau(features)$) to the output variables ($\tau(labels)$). If we have some prior knowledge about the nature/behavior of the data, more strict assumptions can be made when estimating the function, resulting in a *parametric algorithm* [23]. Parametric algorithms have an established functional form for the mapping function and a fixed set of *parameters* (i.e. intrinsic values that define the model and can be estimated from data). Linear Discriminant Analysis (LDA) and perceptrons are some examples of parametric ML algorithms. When we have little to no prior knowledge, *non-parametric algorithms* usually are a better option, as they allow more freedom to learn any functional form from the train data [11]. These algorithms have a variable number of parameters (as the model shifts its form to fit the data) and only assume that samples that are close to one another are likely to have a similar output. Examples of non-parametric algorithms are K-nearest neighbors and Support Vector Machines (SVM).

The actual process of model training, both in parametric and non-parametric algorithms, corresponds to estimating the model parameters (θ) that result in the best achievable fit to the train data. A very important aspect is how the model interprets the concept of performance of the fit, in order to optimize its classification. To do so, certain metrics need to be defined to measure the performance of the model - *performance metrics*. Many of these metrics are derived from a *confusion matrix*, which is a visualization tool that describes the performance of classification. Table 2.3 represents the structure of a confusion matrix for a binary classification problem (i.e. deciding between two possible classes), but it can be extended to multi-class classification problems (i.e. when more than two possible classes exist). The labels ”positive” and ”negative” are usually used in binary classification problems to describe the pattern we wish to detect and the background, respectively. However, other terms can be used to describe each class (e.g. ictal as the positive class, and non-ictal/background as the negative class), particularly in multi-class classification problems, where the binary denomination does not suffice.

Table 2.3: Structure of a confusion matrix for a binary classification problem.

	True Positive Class	**True Negative Class**
Predicted Positive Class	true positive (tp)	false negative (fn)
Predicted Negative Class	false positive (fp)	true negative (tn)

From this matrix, several metrics can be derived and Table 2.4 describes some of the most used ones in classification tasks (the bottom five correspond to the extension of the metrics to multi-class classifications). Each metric evaluates different characteristics of the model, therefore, when choosing a metric to use as a measure of the fit of the model, the choice must be made based on the task ahead.

[11] J. Brownlee. ”Parametric and Nonparametric Machine Learning Algorithms”. [Online].

Moreover, performance metrics can also be used in other situations, such as to evaluate the model on the test data or to provide insight on the behavior of our model.

Accuracy is the most commonly used metric, as it provides a general evaluation of the model (how well it classifies samples in general), however, it assumes equal costs for misclassifications and it performs poorly when there is significant class imbalance because it will always favor the majority class. In some contexts, it is preferable if all instances of an event are found even if that results in misclassifying some negative cases (e.g. when detecting medical conditions, it is better to have a few false alerts than overlooking a true positive) - high sensitivity/recall and low precision. In other contexts, we would rather sacrifice some positive instances, to guarantee that only the positive ones are detected (e.g. in high security access applications, false acceptances are worse than false rejections) - low sensitivity/recall and high precision. This trade-off can be difficult to navigate when the applications are not so obvious, in which cases, the F1-score might be an adequate solution, as it provides a trade-off between the two metrics (see Table 2.4 for more examples of alternative metrics).

In practice, many more metrics can be used as performance measures in classification tasks, but they will not be addresses in the scope of this work.

Generalization, Overfitting and Underfitting

In general terms, one would want a ML model to be able to accurately predict the outcome of unseen instances, based on the knowledge of known instances, and this property is known as *generalization* [21]. However, this is not always possible, either due to the underlying nature of the data, presence of noise or missing information.

Here there is a delicate balance between incorporating all the noise present in the train set, and loosening its fit too much. In the first case, by incorporating all the noise, the algorithm will achieve maximum performance on the instances from the train set, but then it will not be able to generalize to unseen data, since the model does not actually correspond to reality - which is called overfitting (Figure 2.5 C). In the second case, by loosening the fit to the test data too much, it will have a poor performance in both the train set and test set, because it can not capture any of the underlying structure of the data - which is called underfitting (Figure 2.5 A). Therefore, when training a ML model, a careful trade-off is needed between relaxing the fit and achieving high training performances, and it is often achieved through *regularization techniques*.

In Figure 2.5, the appropriateness of the fit is very clear, however, visualization is not always possible (e.g. due to the dimensionality of the data), so it might be a challenge to understand if a model is overfitting to the train data. In these cases, *cross validation* acts as an invaluable tool. There are several cross validation techniques (such as K-fold cross validation, leave-one-out cross validation and stratified K-fold cross validation) but the principle behind them remains the same (as illustrated in Figure 2.6): 1) The train dataset is partitioned into a number of subsets, 2) in each iteration, a subset is hold out (validation set) and the model is trained with the remaining subsets, 3) the model is tested on the held out subset and 4) the process is repeated for each subset and average performance is computed.

By training and testing the algorithm several times with (slightly) different subsets, we get a less

Table 2.4: Performance metrics for classification tasks [24]. Note: tp_i - samples belonging to class C_i correctly classified; fn_i - samples belonging to class C_i incorrectly classified; fp_i - samples not belonging to class C_i classified as such; tn_i - samples not belonging to class C_i not classified as such.

Metrics	Formula	Description
Accuracy (Acc)	$\frac{tp+tn}{tp+fp+tn+fn}$	Ratio of correct predictions over the total number of samples evaluated. Highly sensitive to class imbalance.
False Positive Rate (FPR)	$\frac{fp}{fp+tn}$	Also known as False Detection Rate. Ratio of samples incorrectly classified as positive.
Sensitivity (Sn)	$\frac{tp}{tp+fn}$	Also known as True Positive Rate and Recall. Ratio of positive samples that are correctly classified. Conveys how well the algorithm detects positive instances.
Specificity (Sp)	$\frac{tn}{tn+fp}$	Also known as True Negative Rate. Fraction of negative samples that are correctly classified. Coveys how well the algorithm detects negative instances.
Precision (P)	$\frac{tp}{tp+fp}$	Conveys how well the algorithm detects only the relevant (positive) samples.
F1-score	$\frac{2*\mathsf{P}*\mathsf{Sn}}{\mathsf{P}+\mathsf{Sn}}$	Harmonic mean between sensitivity and precision. Ranges between 0 and 1. It conveys how precise and robust the classifier is and can be seen as a trade-off between sensitivity and precision.
Averaged Accuracy	$\frac{1}{N}\sum_{i=1}^{N}\frac{tp_i+tn_i}{tp_i+fn_i+tn_i+fn_i}$	Average accuracy across all classes.
Averaged Error Rate	$\frac{1}{N}\sum_{i=1}^{N}\frac{fp_i+fn_i}{tp_i+fn_i+tn_i+fn_i}$	Average error rate across all classes.
Averaged Precision (P_{avg})	$\frac{1}{N}\sum_{i=1}^{N}\frac{tp_i}{tp_i+fp_i}$	Average of per-class precision.
Averaged Sensitivity (Sn_{avg})	$\frac{1}{N}\sum_{i=1}^{N}\frac{tp_i}{tp_i+fn_i}$	Average of per-class sensitivity.
Averaged F1-Score	$\frac{2*\mathsf{P}_{avg}*\mathsf{Sn}_{avg}}{\mathsf{P}_{avg}+\mathsf{Sn}_{avg}}$	Average of per-class F1-score.

biased estimation of the performance of our model, and therefore understand if the model has an appropriate fit.

2.2.3 Machine Learning in the Detection and Prediction of Epileptic Seizures

Since the characteristic patterns of each epilepsy syndrome and seizure type are considerably well defined, it is only natural that ML approaches try to create models that approximate these patterns and behaviors of the signal over time. These models, representative of each possible state (i.e. non-epileptic, pre-ictal, ictal and interictal), then allow for the algorithms to compare the newly collected data with the models, and classify them accordingly. However, since there is a considerable variability of the

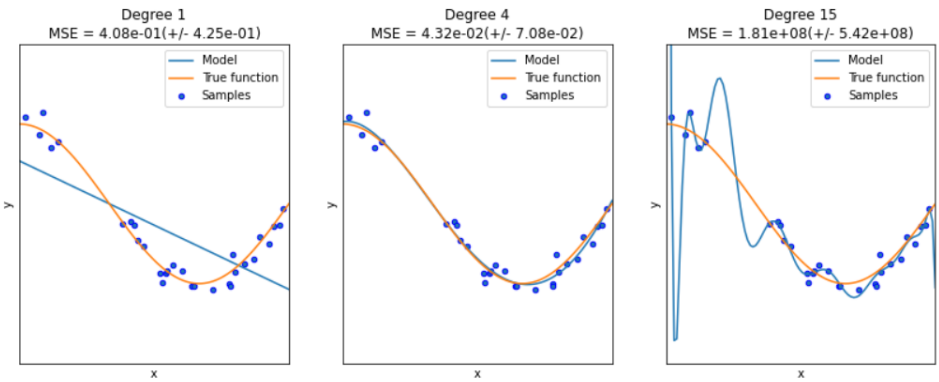


Figure 2.5: Example of Minimum Squared Error (MSE) algorithm, using a polynomial regression with different degrees, which results in (A) underfitting, (B) appropriate fit and (C) overfitting. Reprinted from M. Tripathi. "Underfitting and Overfitting in machine Learning". [Online].

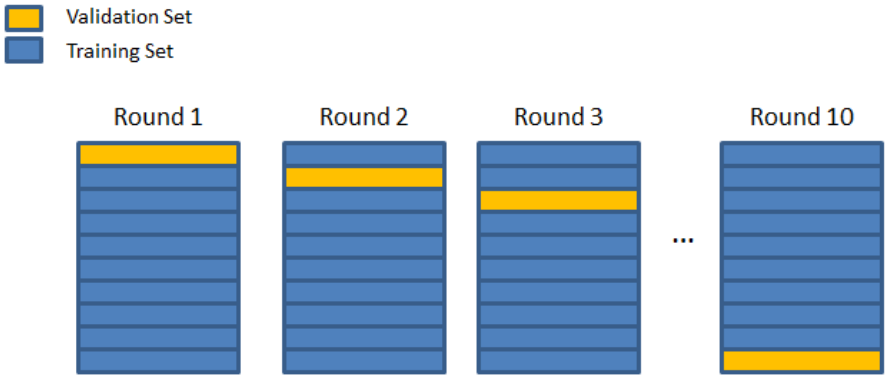


Figure 2.6: Process of cross validation in general. (Source: R. Shaikh, 2018. "Cross validation explained: Evaluating estimator performance". [Online].).

characteristic patterns within each syndrome and seizure type, discriminating them when implementing a ML method may be useful to more accurately define each state and thus increase the accuracy of seizure detection.

In such a system, the two main considerations are: which features to extract from the input signal (analysis of the signal), and the method by which classification is performed. Regarding signal analysis (i.e. feature extraction), EEG is commonly characterized by its spectral features, i.e. the five frequency bands of EEG signals. However, these provide only a small fraction of what constitutes the signal and, therefore, much valuable information is lost if no more features are accounted for. Hence, the analysis of EEG can be performed through various methods that, according to Acharya et al. [16], fall under four main categories: time domain, frequency domain, time-frequency domain and nonlinear methods, each of which provide different information to describe the signal.

The next section provides some background knowledge on one of the methods that will be used in the context of this work for the extraction of features.

Discrete Wavelet Transform

Wavelet Transform (WT) is an analysis method that belongs to the time-frequency domain: contrary to Fourier analysis, which provides spectral information of the signal averaged across its duration (providing no temporal resolution), wavelet analysis allows for simultaneous local evaluation of the spectral and time domains of a signal [17], albeit with different resolutions (high frequency resolution at low frequencies and high time resolution at high frequencies [25]).

This analysis method is extremely powerful, not only for feature extraction but also for denoising, as it allows to decompose the signal in frequency sub-bands and then proceed to its reconstruction (assuming the use of a proper orthogonal basis for the decomposition) [25].

This multi-resolution analysis relies on the use of a mother wavelet that is *scaled* several times. By shifting it across the signal, we obtain the correlation of the signal and that specific scaled version of the mother wavelet. Several mathematical representations of families of wavelets have been proposed, such as Haar, Daubechies, Coiflets and Symlets, and depending on the type of signal, some are more efficient for particular applications (although there are no set rules for this). The mother wavelet is usually chosen to resemble the transients/patterns commonly present in the signal of interest, so that the scaled version of the wavelet has a high correlation with this part of the signal and a low correlations with regions that are not of interest (e.g. noise) [26].

DWT, in particular, has a filter-bank implementation of the WT: it achieves the decomposition by successively applying Finite Impulse Response (FIR) lowpass and highpass filters, followed by a down-sampling of factor 2 (filter bank). At each level, the signal is decomposed in detail coefficients, *cD*, (high-frequency components) and approximation coefficients, *cA*, (low-frequency components, further decomposed afterwards) [27].

Hence, the decomposition of the signal in sub-bands results in a set of coefficients per sub-band, each representing the content of the signal corresponding to a specific range of frequencies. Moreover, the range of frequencies in each sub-band is dependent on the number of decomposition levels and the original sampling frequency, as illustrated in Figure 2.7.

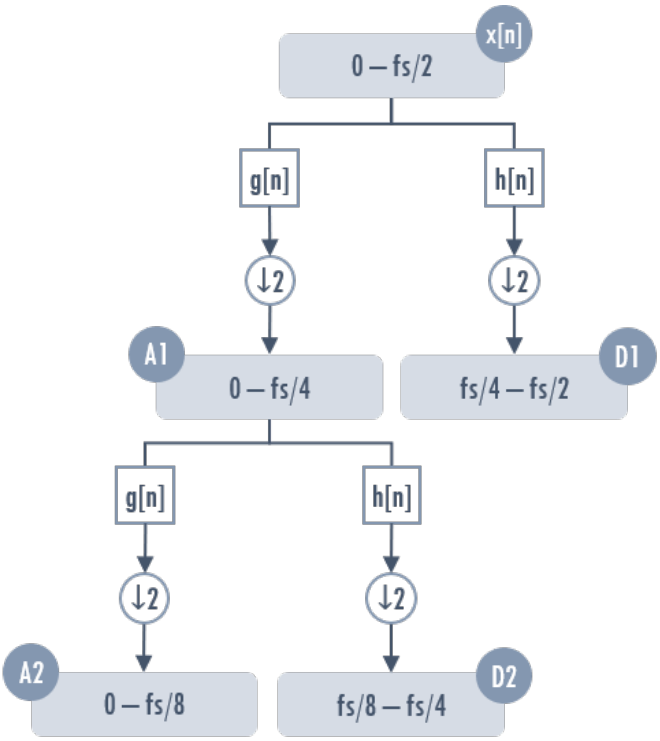


Figure 2.7: Illustration of the decomposition steps in the DWT multilevel decomposition. x[n]: original signal, h[n]: highpass filter, g[n]: lowpass filter, ↓2: downsampling by a factor of 2. Inspired by [25].

2.3 Communication Protocols

The implementation of the automated long-term data acquisition system uses two main technologies, for which background details are also important to provide. In particular the communication protocols that will be implemented on the acquisition setup proposed in Section 4.

2.3.1 Bluetooth

Bluetooth is a standard communication protocol, used to support high-speed, low-powered, wireless technology. It operates in the 2.4 GHz ISM spectrum band (2400 to 2483.5 MHz)[12] and provides real-time data streaming between devices at a short range (dependent on external factors, such as the presence of attenuators) [28].

At the moment, Bluetooth has two different versions: Bluetooth Low Energy (BLE) and Basic Rate/ Enhanced Data Rate (BR/EDR or Bluetooth Classic), each designed for different use cases. The choice of which technology to use highly depends on two key aspects: power consumption and data transfer rate. As the name indicates, BLE is designed for very low power operations, transmitting data over 40 channels within the 2.4GHz ISM Band, at transfer rates ranging from 125 Kbps to 2 Mbps. Classic Bluetooth, on the other hand, still has low power consumption (just not as low as BLE) and performs the transmission of data over 79 channels, at transfer rates from 1 Mbps to 3 Mbps[13].

[12]The frequency band used by Bluetooth has "been set aside by international agreement for the use of industrial, scientific and medical devices (ISM)" [28]

[13]Bluetooth SIG Inc. "Learn About Bluetooth: Radio Versions". [Online].

Communication between devices is guaranteed through a private network called *piconet*. Each piconet is comprised of up-to 8 devices, following a master/slave model: a single device, called *master*, coordinates communication and traffic with the remaining devices, called *slaves*, which are uniquely connected to the master. The actual process of connection (also called forming a *link*) is based on the principle of ”inquiry” and ”inquiry scan”: one of the devices actively sends out an inquiry for connection and nearby scanning devices (which are listening on known frequencies for these inquiries) respond to the inquiry with identifying information (e.g. MAC address[14]). When the user selects the wanted device, connection is established.

In order to avoid interference and enhance security of the communication, after a slave forms a link with the master, a frequency hopping model is established: multiple times per second, all devices in that piconet synchronously change channels together, according to a pattern set by the master[15]. More recent technology was introduced in this system, Adaptive Frequency Hopping, so that the master looks only for available channels, excluding the portion of frequencies already taken by interfering devices (particularly useful when static ISM systems are present, such as Wi-Fi) [16].

2.3.2 MQTT

Message Queue Telemetry Transport (MQTT) [29] is a standard messaging protocol characterized by being lightweight and allowing for bi-directional communication. It is based on device-to-broker and broker-to-device messaging, allowing for multiple devices to be involved, acting as a sort of broadcasting mechanism (architecture illustrated in Figure 2.8). The MQTT broker is the server responsible for receiving the incoming messages from the publishers and routing them towards the proper destination, i.e. the subscribers. This destination is set by a *topic*, which corresponds to a UTF-8 string that can be seen as a ”path” used by the broker to filter messages to the subset of clients subscribed, broadcasting only to that specific subset.

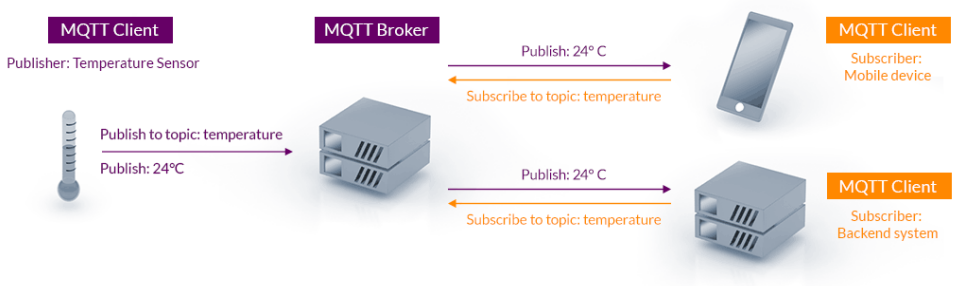


Figure 2.8: MQTT: publish/subscribe architecture, illustrated by [29].

An MQTT session is comprised of four stages: 1) Connection, 2) Authentication, 3) Communication and 4) Termination. Connection is initiated by the client, which establishes a Transmission Control Protocol/Internet Protocol (TCP/IP) connection with the broker through a pre-defined port. Regarding authentication, MQTT provides the option to protect the connection by a username and password (al-

[14]A Bluetooth MAC address is a 48-bit string that acts as a unique identifier of a Bluetooth device.
[15]Honeywell, 2018. ”What is Bluetooth Adaptive frequency Hopping (AFH)”. [Online]
[16]M. Foley. ”How does Bluetooth work”, 2007. [Online].

though not mandatory, as it might not be desirable in some use cases), which is used as a security layer, rejecting unwanted connections. During the communication process, a client can either *subscribe* to a topic (listening to all incoming messages), *unsubscribe* or *publish* to a topic (i.e. actively sending a message to all subscribers, including itself if subscribed). Each message is sent in a byte array (packet) with a format as shown in Figure 2.9, restricting it to a rather small code footprints, hence its lighweight nature [17]. When the client wants to terminate the MQTT session, it sends a message to the broker and closes the connection.

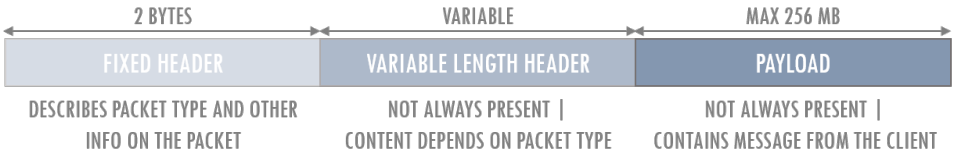


Figure 2.9: MQTT standard packet structure. Content from [30].

Compared to other communication protocols, such as the widely used HTTP, MQTT has an appealing feature, as it ensures high delivery guarantees and provides customization of delivery service quality levels: 0 - at most once, 1 - at least once, 2 - exactly once [29]. As such, it is optimized for high-latency or unreliable networks.

[17] M.Rouse, A. Gillis and P. Waher. ”MQTT (MQ Telemetry Transport)”. [Online].

Chapter 3

State of the Art

Automated EEG analysis for the detection of epileptic seizures is extensively studied in literature, encompassing several feature extraction techniques and machine learning methods, many of which show satisfactory performances. However, high-density Electroencephalography (EEG) acquisition is only reasonable in clinical settings, whereby a section in this chapter is reserved to the research of EEG configurations with a limited number of channels. Beyond EEG, there are several other modalities that are being studied in the scope of seizure detection and it is interesting to keep in mind how these evolve alongside the modality that was chosen as the main focus of this work. The end of this chapter is dedicated to the state of the art of wearable solutions for the acquisition of EEG, since it is an essential complement to this work.

3.1 Tools for Seizure Management in Epilepsy

The main concept behind seizure management is seizure tracking, which not only allows to better understand and categorize the condition, but also to recognize seizure triggers in order to avoid them and thus reduce their frequency. There are two major categories of tools that address this need: seizure diaries and seizure monitoring.

Seizure diaries allow patients to perform self-monitoring by logging seizure events and environmental factors that could have contributed to the occurrence or aggravation of the seizure. Several seizure diary tools can be found online, either in paper form (e.g. Epilepsy Action and Epilepsy Society) or online/mobile app form, such as the ones provided by Epilepsy Ireland and Epilepsy Foundation. Table 3.1 describes the key features that paper-form and mobile app seizure diaries usually provide. There are several tools available, particularly mobile apps, however, they all provide roughly the same features.

Seizure monitoring, on the other hand, is usually performed with resort to video-EEG in which the patient is subject to continuous observation. Both EEG and video are recorded for long periods of time, with the purpose of recording seizures and characterize the condition. Evidently, this setup is usually achieved through restrictive conditions, either with hospitalization of the patient or in ambulatory settings, but requiring a bulky EEG head cap and video equipment.

Table 3.1: Comparison of features commonly provided by paper-form and mobile app seizure diaries [31].

	Paper form	**Mobile App**
Monthly seizure overview	✓	✓
Description of seizure occurrence	✓	✓
Possible triggers	✓	✓
Medication side-effects	✓	✓
Medication changes	✓	✓
Video recording		✓
Emergency action		✓
Visual reports/trends		✓
Share electronically		✓
Possibility of data sharing for research		✓

With the growth of Machine Learning (ML) and wearable technology, the possibilities of combining these two tools in a single system, capable of long-term monitoring, event detection and recording of environmental factors, becomes palpable.

3.2 Electroencephalography-Based Automatic Seizure Detection in Epilepsy

Due to its relation with cerebral activity (see Section 2.2.1), the EEG is the first approach when studying brain-related conditions, such as epilepsy. Naturally, this biosignal modality has been the standard in the scope of epilepsy, having been widely explored in previous research and thus providing much information about the extent of its detection and prediction capabilities. The appeal of EEG greatly stands on the fact that this biosignal features characteristic patterns seen only in an epileptic individual, even during non-ictal periods, which allows to diagnose the condition, and presents other characteristic patterns before and during ictal periods, allowing for the detection and prediction of epileptic seizures [16].

Epileptic seizure detection in clinical settings is most commonly performed by expert physicians who analyze hours-long EEG recordings by visual inspection and manually annotate the beginning of the seizures (distinguishing between ictal, pre-ictal and inter-ictal segments). This is obviously a time-consuming, tedious and error-prone approach. The correct identification of the seizures is crucial to efficiently administrate drug therapy and, therefore, reduce the risk of future seizures and complications [32]. Hence, automatic seizure detection in clinical practice provides an invaluable aid.

There is extensive literature on the algorithms used for automatic detection of epileptic seizures; in the next section, we will focus on the features that may be used as input for these algorithms (disregarding the number of channels used) however, from that section on, we will focus on studies that provide some insight on wearable systems for the task of seizure detection.

3.2.1 Features for Seizure Detection

Seizures affect most of the EEG frequencies, hence why many studies have explored frequency-based features (i.e. spectral analysis) for the detection of epileptic seizures. Birjandtalab et al. [18]

extracted this spectral information by calculating the normalized PSD of each frequency band, achieving a sensitivity of 80.9% and precision of 47.4% in the CHB-MIT database. Similarly, Fergus et al. [33] transformed the time domain signals using PSD and from that extracted features such as frequency parameters (e.g. peak frequency), measures of the complexity of the signal (e.g. SampEn) and statistical parameters (e.g. skewness and kurtosis). The authors reported a sensitivity and specificity of 88%, also in the CHB-MIT database

Since EEG contains non-stationary characteristics, as an alternative to purely spectral features, time-frequency methods have been frequently used in feature extraction in machine learning problems, and the field of epilepsy detection is no exception to this [34]. Particularly, Discrete Wavelet Transform (DWT) allows to decompose the signals into frequency sub-bands, given by sets of coefficients [35]. This sub-band extraction process comprises, in many studies, the whole feature extraction task.

However, DWT decomposition has also been used as a first step in the process of feature extraction, being followed by some other extraction procedure, such as average power, mean, entropy, standard deviation relative wavelet energy [25] and line length (measure of signal complexity) [36].

The nonlinearity of EEG has been studied in literature (Mirzaei et al. [37], Pradhan et al. [38]) and it arises from the innate nonlinear behavior of biological mechanisms, such as the threshold-dependent firing of neurons, that results in a signal made up of many sinusoidal components of different frequencies. Although linear systems are more robust and more intuitive to interpret, nonlinear methods, such as Approximate Entropy (ApEn), Sample Entropy (SampEn), Fractal Dimension (FD) and Hurst Exponent (HE), present an indisputable advantage: their ability to capture the nonlinearity in the behavior of biosignals [39].

At the time of writing, the state-of-the-art reveals no study aiming at discriminating the best features towards seizure detection with a limited number of EEG channels. However, there have been some attempts towards identifying the best performing features in general terms, which is a great starting point. Logesparan et al. [40] analyzed the set of 65 features, all previously used in publications regarding this thematic, enumerated in Table 3.2. The predictive power of each feature was evaluated individually with a simple threshold-based algorithm. The authors reported the highest overall performances with line length - with a sensitivity of 85.1% and specificity of 92.2% - and DWT-based features, reporting the best overall sensitivity of 89.4% as well a specificity of 92.9%.

Table 3.2: Comprehensive set of features analyzed in [40], in terms of their predictive power in seizure detection. (*): computed across 4 frequency ranges - $D3$ (12.5-25 Hz), $D4$ (6.25-12.5 Hz), $D5$ (3.125-6-25 Hz) and $A5$ (0-3.125 Hz).

Time Domain (TD)	Complexity, energy/power, fractal dimension, kurtosis, line length, maximum, mean, minimum, mobility, non-linear energy, relative derivative, Shannon entropy, skewness, total maxima and minima, variance/standard deviation, zero crossing, zero crossing of first derivative.
Fourier Transform (FT)	Median frequency, peak frequency, power*, spectral edge frequency, spectral entropy*, total spectral power.
Continuous Wavelet Transform (CWT)	Coefficient z-score, energy, entropy, standard deviation of energy.
Discrete Wavelet Transform (DWT)	Bounded variation*, coefficients*, energy*, entropy*, relative bounded variation*, relative power*, relative scale energy*, variance/standard deviation*.

3.2.2 Limited-Channel Electroencephalography Seizure Detection

High-density EEG is only achievable in a clinical setting, particularly because it is uncomfortable and stigmatizing, since it requires the use of a head cap full of electrodes [18]. Moreover, applying ML algorithms to a large number of channels (e.g. 23 channels, from the commonly used TCP montage) is computationally expensive, since each sample is represented by $\#features \times \#channels$. This results in a significant burden, both during the process of feature extraction and during the actual algorithm implementation. Besides, aiming at the implementation of real-time seizure detection systems, this additional burden can significantly contribute to the decrease of battery life of the hardware systems, which is definitely not desirable. Moreover, the long-term use of clinical scalp EEG setups is seen as undesirable by many epileptic patients [41]. As such, a paradigm shift is needed towards wearable devices for EEG acquisition.

There is some literature addressing the channels/montages that might be more significant for the task of seizure detection, using varied techniques that often shy away from the demanding exhaustive search of all combinations of channels. These studies commonly use clinical EEG recordings and select the desired channels from those, which is obviously not the same as using wearables for the acquisition, in terms of noise, signal power and localization of electrodes (to some extent). However these studies are crucial for the optimal design of wearables.

Birjandtalab et al. [18] employed a *filter technique*, based on Random Forest (RF), with the purpose of finding the most relevant/informative channels and filtering out the remaining. In this study, the authors extracted a set of 5 spectral features for all channels of the TCP montage and used the sum of impurities as the splitting criteria. The channels were sorted according to their contribution in the forest (i.e. how many times a feature from the ith channel appeared in the forest). The authors reported a significant difference in contribution from the channels, placing channels 17, 6 and 8 as the most informative. However, they only reported the findings for one of the patients and failed to mention the correspondence of these to the TCP montage denomination. To validate their findings, the authors performed classification using the limited set of 3 channels (specific to each patient) and achieved an average sensitivity of 80.9% and precision of 40.5%.

Moctezuma and Molinas [42] recently experimented a different approach, using both Backward-Elimination and a Genetic Algorithm (GA) to identify the channels that provided the best seizure detection performance. This analysis was performed independently for each patient and the performance metric used was accuracy. The authors reported very high performances (93.5 ± 6.0 and 95.2 ± 5.3, across all patients) for a small number of channels (1 and 2, respectively), in some cases using Backward-Elimination and others using GA. However, using accuracy as the sole performance measure can be deceptive: in epilepsy datasets, there is a large imbalance between the number of segments belonging to non-ictal and ictal categories; hence, a large accuracy can correspond to classifying every sample as non-ictal, disregarding ictal segments, unless some balancing measures are taken. The authors did not report any clarifying measures.

In 2011, Peterson et al. [43] suggested an interesting approach for the detection of absence seizures, in particular, investigating the detection performance of single channels, from a set of 18 channels. The study was performed on 19 patients with childhood absence epilepsy, extracting the energy of WT sub-bands and using SVM as the classifier, coupled to a temporal constraint on seizure identification (i.e. samples were only classified as ictal, as long as a minimum of 2 consecutive seconds are classified as such). Their findings indicate that frontal channels are overall better at discriminating this type of seizures, as illustrated in Figure 3.1. The authors reported that the best overall channel was *F7-FP1* (with a sensitivity of 99.1% and a FPR of 0.5/h), followed by *F7-F3* (with the same sensitivity but with a FPR of 1.0/h). *Fp1-Fp2* (which is particularly relevant for this work) also achieved a reasonable performance (with a sensitivity of 93.7% and FPR of 1.4/h).

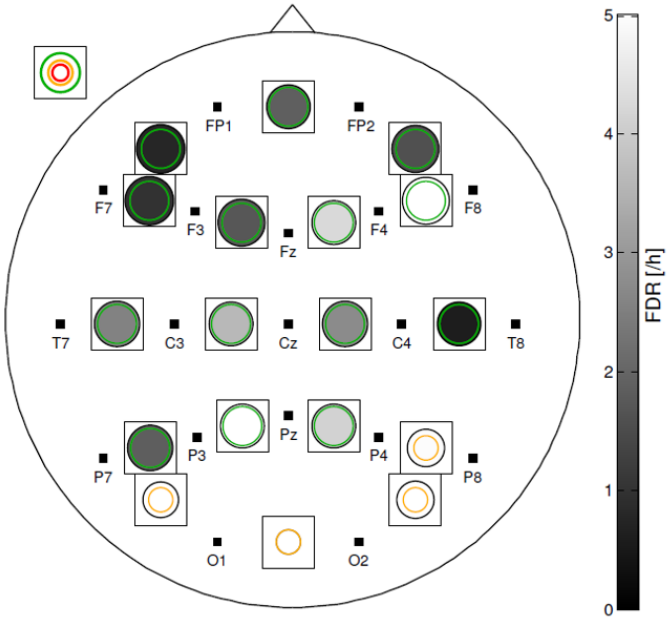


Figure 3.1: Topographical distribution of the overall performance of all channels (located between the two electrodes that generate the specific channel). Sensitivity is given by the diameter of the corresponding circle (red: 50%, yellow: 70% and green: 90%). FPR is given by the color, as shown in the colorbar on the right. Source: Peterson et al. (2011) [43].

Lin et al. [44] investigated the detection performance of the channel *Fp2-F8* for generalized seizures (on 7 subjects from the CHB-MIT database), with the aim of designing a wearable EEG acquisition system. The detection algorithm used ApEn as the discriminant feature and an LDA classifier. The authors reported an average detection rate of 92.7% and FPR of 0.527/h. Similarly, Sopic et al. [45] proposed a wearable, in the format of glasses, which acquires signal corresponding to the channels *F8-T4* and *F7-T3*. The authors validated their system with the CHB-MIT database (10 subjects), comparing the use of this set of channels with all the available channels of the dataset. They reported comparable results achieving an average sensitivity of 93.8% vs 97.0% (all channels) and an average specificity of 93.4% vs 95.8%.

Shah et al. [46] also tested a set of channels for seizure detection (not specifying a type in particular), basing their configurations (6 in total, ranging from 2 to 22 channels) on domain knowledge.

The authors used a hybrid CNN-LSTM deep learning approach, reporting relatively low average performances: achieving the best performance with the 22-channel configuration (with a sensitivity of 39.2% and specificity of 90.4%), producing relatively similar results with the 20, 16 and 8-channels configurations and reporting the worst performance with the 2-channel configuration (with a sensitivity of 31.2% and specificity of 40.8%). However, some studies report that deep learning has not yet shown consistent improvements over the commonly used methods, when it comes to EEG data, which could perhaps explain such low performances.

Table 3.3: Summary of the methods and results from the articles described in this section. (*): patient-specific approach, (s.r.): self-recorded. PSD: Power Spectral Density, FD: Fractal Dimension, ApEn: Approximate Entropy, LFCCs: Linear Frequency Cepstral Coefficient. KNN: K-Nearest Neighbors, SVM: Support Vector Machines, NB: Naive Bayes, RF: Random Forest, LDA: Linear Discriminant Analysis, CNN: Convolutional Neural Network, LSTM: Long Short-Term Memory Network. Acc: Accuracy, Sn: Sensitivity, Sp: Specificity, P: Precision, FDR: False Detection Rate.

Article	Dataset	Features	Classifier	Channels	Performance
[18] (*)	CHB-MIT	PSD of sub-bands	KNN	3 channels	Sn: 80.9%, P: 40.5%
[42] (*)	CHB-MIT	Sub-band energies + FD	SVM + KNN + NB + RF	1 channel 2 channels	Acc: 93.5% Acc: 95.2%
[43]	(s.r.)	Sub-band log-sum energies	SVM	F7-FP1 F7-F3 Fp1-Fp2	Sn: 99.1%, FDR: 0.5/h Sn: 99.1%, FDR: 1.0 Sn: 93.7%, FDR: 1.4/h
[44]	CHB-MIT	PSD of sub-bands + ApEn	LDA	Fp2-F8	Sn: 92.7%, FDR: 0.527/h
[45]	CHB-MIT	PSD of sub-bands + entropies	RF	F8-T4, F7-T3	Sn: 93.8%, Sp: 93.4%
[46]	TUSZ	LFCCs	CNN-LSTM	P4-O2, C3-Cz	Sn: 31.2%, Sp: 40.82%

Overall, these findings seem to suggest the possibility of performing seizure detection in a non-clinical environment with the use of practical wearables (with very few electrodes). However, it is important to note that both [18] and [42] performed their analysis for each patient individually, failing to evaluate the feasibility using a predefined limited-set of channels, and thus relying on flexible positioning of the electrodes of the wearable and a period of training for every new patient (which might not always be reasonable).

3.3 Use of Other Modalities for Epileptic Seizure Detection

Although several wearable solutions for the acquisition of EEG have been proposed (as will be described in Section 3.4), most are still in development phase, lacking in acceptance and/or validation across the clinical community. As shown in the previous section, there is a large potential for the emergence of detection algorithms based on a limited number of EEG channels, that can be transposed to the design of patient-acceptable wearables. However, in parallel to the evolution of this modality, researches have been trying to take advantage of the immense boost in market acceptance of wearable healthcare technology[1] (from fitness trackers to medical-grade wrist-ECG), by incorporating peripheral physiological signals to the field of Epilepsy [5].

While for tonic-clonic seizures (characterized by violent muscle contractions) motion-based methods, such as surface Electromyography (sEMG) and Accelerometer (ACC), provide all the necessary inputs for the detection of ongoing seizures, they fail to provide relevant information for most of the remaining seizure types or for the task of seizure prediction. Hence, when using other modalities beyond EEG, a multimodal approach is often desirable to make sure the wide variety of manifestations (particular to each type of seizure) can be accounted for [5]. Commonly, researchers resort to Electrodermal Activity (EDA), respiration, Electrocardiography (ECG) and Photoplethysmography (PPG) [47].

One such example of a multimodal approach to seizure detection, namely with the use of EDA and ACC, is Empatica's Embrace smartband (however limited to the recognition of ongoing tonic-clonic seizures). It was the first commercially available multimodal wristband (along with Empatica's E4), non-EEG based, physiology-signal seizure monitoring system [48]. This device recognizes ACC and EDA signatures to detect generalized tonic-clonic seizures and alert caregivers. Empatica published a study [48] regarding this device, however only made a made brief reference to previously unpublished results (sensitivity of 92%-100% with an FPR of 0.42/day), failing to provide methodology or other assessment information [49].

On the other hand, currently available literature is particularly rich in studies performed with cardiac-based analysis of Epilepsy. Previous studies have reported that the analysis of Autonomic Nervous System (ANS) activity may provide some hindsight on the detection/prediction of seizures. Heart Rate Variability (HRV), in its turn, provides information on ANS activity [50], hence why a significant number of studies in the field of cardiac-based epilepsy analysis have explored its discriminant properties in the scope of epileptic seizures.

Vandecasteele et al. [51] compared the performance of wearable ECG (180° eMotion Faros) and PPG (Empatica E4 smartwatch) devices with standard ECG in the detection of temporal-lobe epileptic seizures. In this study, the detection was based on HRV (extracted from the ECG) and Pulse Rate Variability (PRV) (extracted from the PPG) and the classification was performed with a SVM classifier, with a Gaussian kernel. The authors reported relatively varied results, with an average sensitivity of 70% for the wearable ECG, which performed better than the standard ECG (which achieved an average sensitivity of 57%) and a sensitivity of only 32% for the wearable PPG. The authors attributed the low

[1] Business Insider. "Latest trends in medical monitoring devices and wearable health technology". [Online].

performance of the last device to an unreasonable number of motion artifacts.

Interestingly, Rukasha et al. [49] assessed 12 primary studies regarding the performance of wearable electronic devices in the detection of epileptic seizures (including [51, 48] cited in this section) and found that a majority of studies aim at detecting seizures associated to motor manifestations (such as generalized tonic-clonic), which seems to suggest this as a limitation in most of the available non-EEG based systems.

Although Shoeb and Guttag [52] did not aim at a wearable, every-day seizure detection algorithm, they studied the effect of complementing EEG with ECG, in a patient-specific detection algorithm. The authors reported an increase in detection from 6/10 seizures to 10/10 and a decrease in FPR from 0.38/h to 0.21/h. Similarly, Sirpal et al. [53] investigated the possibility of integrating EEG with another modality - functional Near-Infrared Spectroscopy (fNIRS). This relatively novel optical technique is used to indirectly measure brain activity, through the dynamics of oxygenated and deoxygenated hemoglobin (which provides the spatial resolution that EEG lacks). With this experiment, the authors reported an increase in sensitivity from 85.2% to 89.7% and in precision from 82.8% to 87.3%. Although non-wearable seizure detection systems are out of the scope of this work, these findings point towards a promising combination of EEG and ECG/fNIRS, which may well transpose to the ”wearable version” of these modalities.

3.4 Wearable Solutions for Electroencephalography Acquisition

Several wearable devices have been developed for seizure detection using other modalities, as the ones featured in Section 3.3, however EEG is the bio-signal with most direct potential for this task [44]. The demand for portable units for both clinical and non-clinical use motivated the introduction of the concept of wearable EEG, abandoning the need for wires, bulky head caps and gels [54].

Casson (2019) in [54] reviewed the EEG wearables available by the end of 2018. All the wearable EEG units featured in the article were wire-free, having wireless transmitters, and had miniature batteries, incorporated within the cap. However, all of them had more than 23 channels, and were therefore incorporated in full head caps. This is a major bottleneck regarding their use in non-clinical, real-life settings, being more suited for ambulatory settings. The author acknowledged the trend of evolution of EEG hardware towards more flexible and socially discrete devices placed on non-haired regions.

Researchers have developed more discrete wearable solutions for the acquisition of EEG, comprising two main architectural designs: behind-the-ear technology, such as mBrain Train’s cEEGrids, and headbands, such as BrainBit’s Smart EEG Headband and EmotAI (which will be discussed throughout this work). More recently, Epitel is developing a more flexible solution that does not match any of the two designs.

cEEGrids is a novel technology based on multi-channel gel sensor arrays that are placed around the ear using an adhesive. Debener et al. (2015) [55] validated this technology in the recording of meaningful continuous EEG, event-related potentials and neural oscillations. Gu et al. (2017) [56] also performed a comparison study regarding conventional and behind-the-ear EEG, reporting comparable

seizure detection performance in patients with focal epilepsy (average sensitivities of 81.2% vs 82.2%, with FPR 1.36/h vs 1.15/h, respectively).

The technology that is being developed by Epitel consists of a single-channel electrode wearable patch-like device, designed for long-term EEG recording. Unlike most devices, it provides the flexibility of configuration around the scalp, which can be used to personalize the acquisition and detection to each patient, according to their seizure patterns [18]. Although it is not yet FDA approved and no validation studies were found, the technology has immense potential in terms of applicability in real-life settings.

Table 3.4: Specifications of the mentioned wearable solutions for the acquisition of EEG.

Device	**Type of sensor**	**Sensors**	**Battery**
mBrainTrain's Smartfones	Semi-dry electrodes (unipolar)	11 channels: 3 central + cEEGrids (4 around each ear)	4 hours
BrainBit's Smart EEG Headband	Dry electrodes (spring-loaded gilded sensors)	4 channels: T3, T4, O1, O2	12 hours
Epitel's Epilog	Dry electrodes	1 channel (flexible positioning)	-
EmotAI	Dry electrodes	2 channels: Fp1, Fp2	10+ hours (extendable)

Chapter 4

EpiBOX - Biosignal Acquisition Setup

4.1 Implementation

The designed setup is composed of three main elements, which interact through different routes/ mechanisms: 1) An autonomous recording unit; 2) A biosignal acquisition system; and 3) A mobile app. The setup is responsible for performing signal acquisition, storing the acquired signal and providing its visualization to the user. Figure 4.1 illustrates the elements involved in the setup and the communication channels between each, which might be useful to revisit throughout this Section.

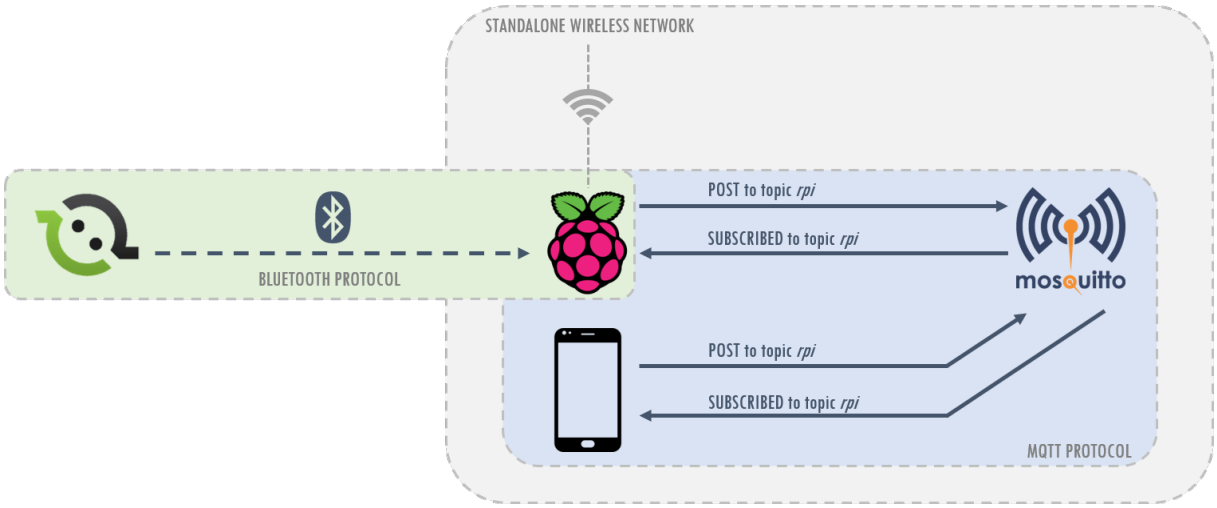


Figure 4.1: Architecture of the proposed system: elements involved and communication channels between each element. The area colored in green corresponds to the communication channel between the Raspberry Pi (RPi) and the acquisition devices (Bluetooth); the area colored in blue corresponds to the communication channel between the RPi and the mobile app (Message Queue Telemetry Transport - MQTT), which is purposefully contained within the grey colored area (wireless network). mosquitto: MQTT broker.

The autonomous recording unit is based on a Raspberry Pi (RPi), which is a tiny, single-board computer that provides a good trade-off between portability and processing power, for this particular application. It acts as the main driver of the whole system, being the sole communication agent with the acquisition devices (described in Section 4.1.1) and being responsible for the launching and mainte-

[1]When the Raspberry Pi is used not connected to a monitor or keyboard, it is known as a headless setup.

nance of the communication channel with the mobile app. The mobile app, fully developed in Flutter, acts as a user-friendly interface with the headless RPi[1], through a pre-defined set of messages that trigger certain actions on the Raspberry Pi, such as which channels to acquire or when to start the acquisition (see Appendix A.4, for the whole set of messages). These messages are exchanged between the two elements through a MQTT protocol, further explained in Section 4.1.2.

MQTT however, is built upon a wireless network, which might not be available in every home/hospital environment. Hence, RPi guarantees the setup of a standalone network, as described in Section 4.1.3.

As in any system, both the hardware and software are not 100% reliable, and in a setup like this one, in which several components and systems are involved, the probability of error occurrence is not negligible. Hence, in order to avoid misrecordings (and consequent waste of resources), it is of utmost importance for the user to be able to guarantee that the acquisition is being correctly performed. The setup provides the user some insight on the acquisition process, by allowing real-time visualization of the signal being acquired. This visualization is performed by streaming the data to the mobile app, through the MQTT channel and the data is plotted as described in Section 4.1.4.

Figure 4.2 illustrates, in general terms, what would be the flow and operation of the acquisition setup in a real environment. Appendix A.2 illustrates how these processes are achieved through the mobile app, including some exemplifying screenshots.

It is important to state that all communication protocols described in the sections below were integrated either with Python, in the case of the RPi, or Flutter, regarding the mobile app.

4.1.1 Biosignal Acquisition System

For data acquisition, this work builds upon previous and ongoing work from the Pattern and Image Analysis (PIA-Lx) research group at Instituto de Telecomunicações (IT), and uses BITalino [57]. BITalino is a hardware and software toolkit, specifically designed to acquire body signals, which has demonstrated to be a valid equipment for research [58], superior to several of its counterparts [59].

This toolkit allows the flexibility of connecting several different sensors beyond EEG (e.g. ECG, EDA, and several others); as well as enabling the reconfiguration of the setup to different physical formats. The device provides the raw signals produced by the Analog-to-Digital Converter (ADC), which can be easily accessed - a crucial feature for the development of *EpiBOX*. Furthermore, it allows for easy interoperability with a range of platforms, including Raspberry Pi and provides an easy-to-use Python module. The connection of the RPi to each BITalino is configured whenever the acquisition process is triggered, through the methods provided in this module.

An invaluable component of this toolkit is the Bluetooth module, that acts as the communication channel between the BITalino and the RPi, allowing for real-time wireless data streaming. With such system, it should be possible to connect up to 7 BITalino devices to the RPi, however, three to four devices is a practical limit, even if the transfer rates allow it[2] (considering, for example, the maximum

[2]Apple Support, "Using a Bluetooth mouse, keyboard or trackpad with your Mac". January 29, 2020, `https://support.apple.com/en-gb/HT201171`. (accessed December 30, 2020).

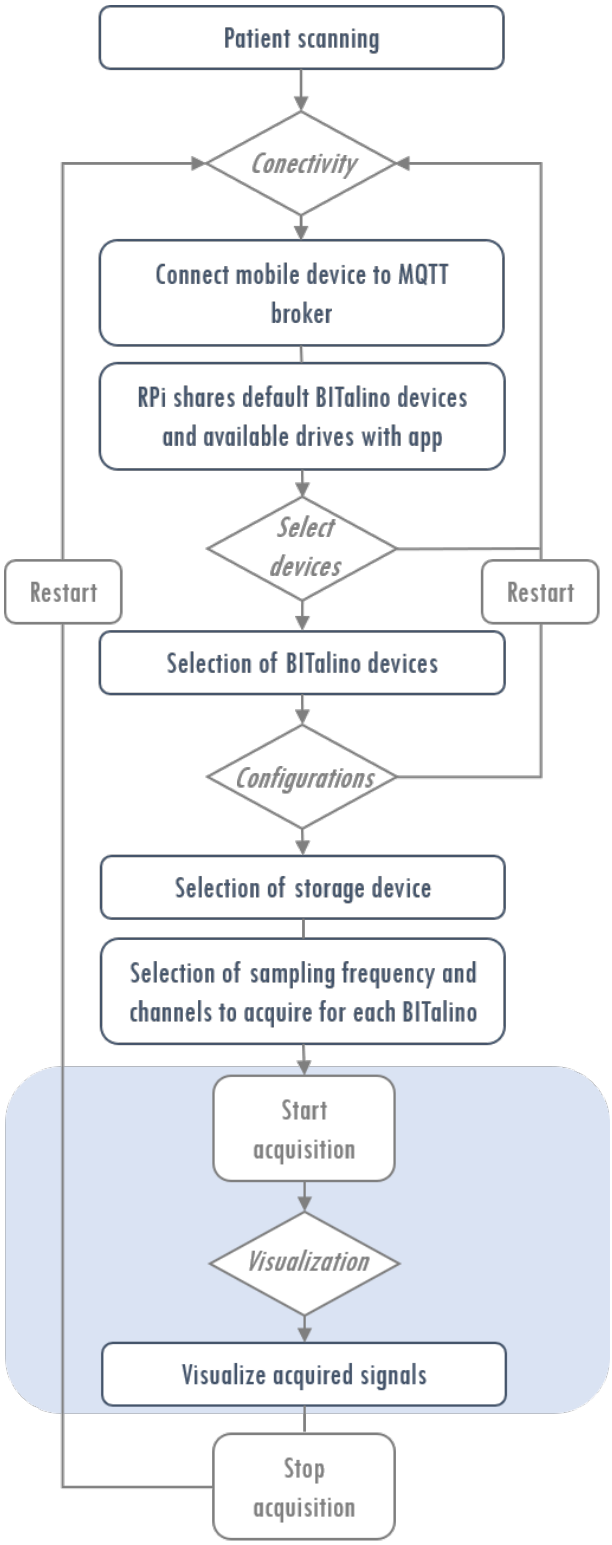


Figure 4.2: Flowchart describing flow and operation of the acquisition setup. Note: the steps highlighted in blue are further illustrated in Figure 4.3.

allowed sampling frequency, i.e. 1 kHz, which would generate a volume of data of 64 Kbps per BITalino[3]). For the purpose of this work, considering that each BITalino allows to acquire up to 6 analog channels simultaneously and synchronously, a maximum of 2 devices were accounted for.

[3]BITalino documentation on the official website defines the reliable throughput of the BT module as 64 Kbps.

4.1.2 Communication Protocol

As stated in Section 2.3.2, MQTT is an extremely lightweight and reliable communication protocol. It is particularly suitable for event-driven machine-to-machine communication, making MQTT very fit for our purpose of communicating between the Raspberry Pi and the mobile phone[4].

There are several open source MQTT brokers available online, with differing capabilities. Although Eclipse Mosquitto does not perform as well in terms of scalability, when compared to other solutions, it was the broker chosen for this work. Scalability is not an issue in this particular application, and this solution does provide easy-to-implement client capabilities both for Python and Flutter, hence the reason for choosing it.

For this work, the RPi was set as the MQTT broker, launching it on boot on the local network. Both the RPi (as a client) and the mobile phone connect to the broker, protected by password, and subscribe to the topic *"rpi"*, ensuring communication between the two, while increasing the privacy of the transaction. This architecture corresponds to the right portion of the illustration of the setup (Figure 4.1).

4.1.3 Standalone Wireless Network

For the communication between the devices to be settled, there needs to be a network available, which is responsible for providing the environment in which the MQTT protocol is established and can create the channel for the message exchange. The Raspberry Pi (RPi) was configured to work as a wireless access point to run a separate, private network. Without access to the Ethernet, the RPi can not provide access to the internet, however, for this particular setup, this acts as an advantageous asset; it provides the means needed to set the MQTT channel within the network, while enhancing the overall security. For an additional layer of security, connection to this network is password protected.

Within this network, the RPi is configured to a static IP address and is responsible for dynamically configuring the IP addresses of the remaining hosts (devices) that connect to the network (in this case, the mobile phone that is running the mobile app).

4.1.4 Data Storage and Visualization

In every acquisition with this setup, there are two major procedures happening: the storing of the acquired data in a *.txt* file (saved with the identification of the patient and the acquisition session) and the real-time plotting of the acquired data whenever the user accesses the mobile app. They happen concurrently with the actual acquisition by the BITalino device(s). Figure 4.3 describes the flow of the process that allows to perform these two procedures.

Conversion to Physical Units

BITalino provides all data in the range 0-1023, corresponding to the 10-bit digital codes produced by the analog-to-digital converter. Hence, to visualize data with the correct physical units, it needs to

[4]The fact that the communication between the Raspberry Pi and the mobile phone is entirely based on rather short strings, further enhances the suitability of MQTT protocol.

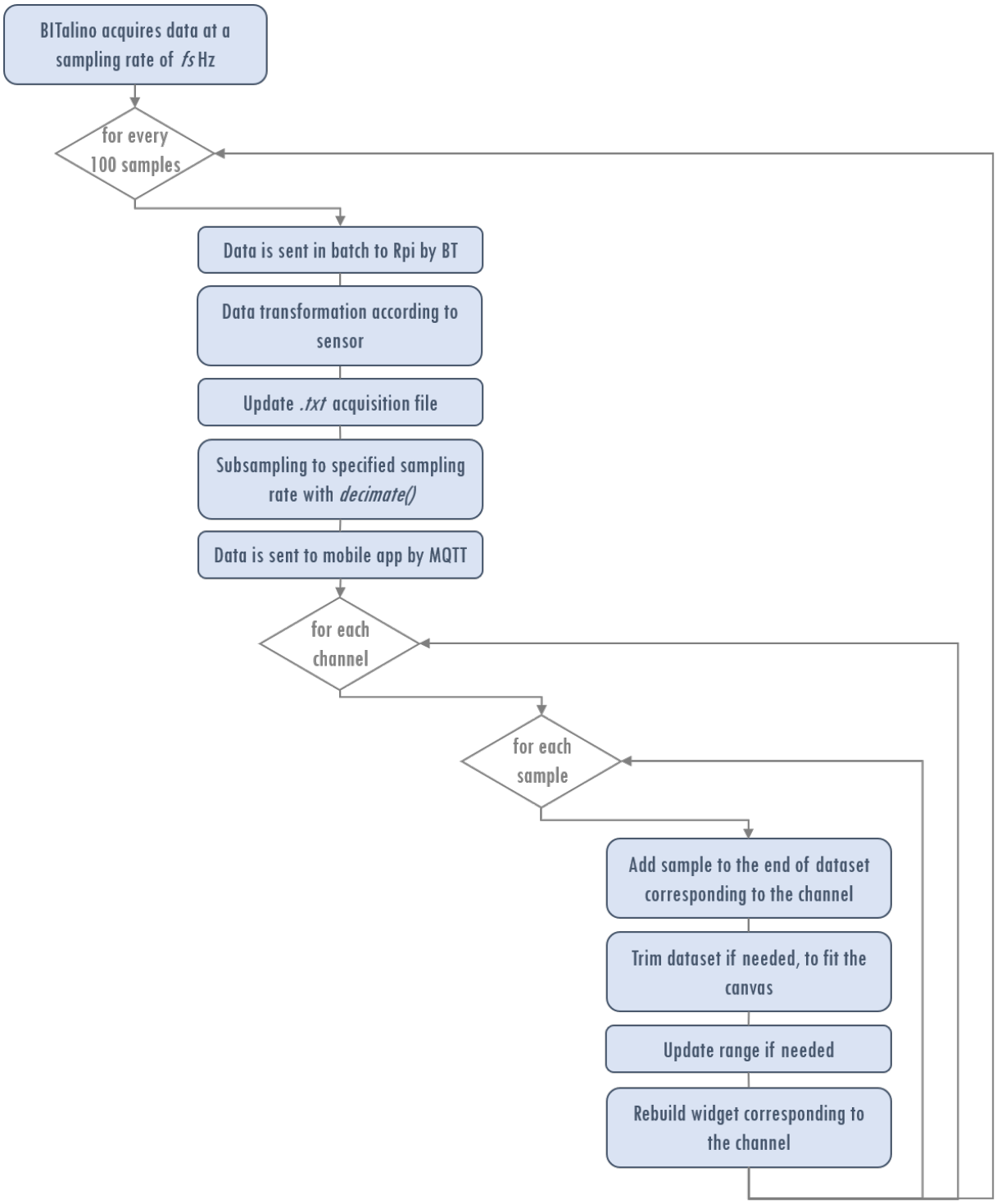


Figure 4.3: Flowchart describing the procedures involved in storing and visualizing the acquired signals. Note: *channel* corresponds to the list of samples acquired in a channel (e.g. channel A1 from one of the BITalinos). *fs*: sampling rate (can be set to [1,10,100,1000] Hz), BT: Bluetooth.

be converted. Each sensor's official datasheet (on BITalino's website) provides a transfer function that allows for the correct conversion (see Table 4.1). Although the mobile app allows for the discrimination of the sensors ACC, PPG and SpO_2, the first requires a calibration (which is not practical to perform) and BITalino's website does not provide the transfer functions for the latter two, hence why they are not converted.

Table 4.1: Transfer functions to convert from ADC to physical units and ranges for each type of biosignal. Given by each sensor's official datasheet, on BITalino website. *ADC*: digital value produced by the analog-to-digital converter; *Vcc*: operating voltage (3.3V); *n*: number of bits of the channel.

Sensor	Transfer Function	Sensor Gain (G)	Range	Unit
EEG	$EEG(\mu V) = (\frac{ADC}{2^n} - \frac{1}{2}).\frac{Vcc}{G}.1.10^6$	41782	$[-39.49, +39.49]$	μV
ECG	$ECG(mV) = (\frac{ADC}{2^n} - \frac{1}{2}).\frac{Vcc}{G}.1000$	1100	$[-1.5, +1.5]$	mV
EOG	$EOG(mV) = (\frac{ADC}{2^n} - \frac{1}{2}).\frac{Vcc}{G}.1000$	2040	$[-0.81, +0.81]$	mV
EMG	$EMG(mV) = (\frac{ADC}{2^n} - \frac{1}{2}).\frac{Vcc}{G}.1000$	1009	$[-1.64, +1.64]$	mV
EDA	$EDA(\mu S) = \frac{ADC}{2^n}.\frac{Vcc}{0.132}$	-	$[0, +25]$	μS
PZT	$PZT(\%) = (\frac{ADC}{2^n} - \frac{1}{2}).100$	-	$[-50, +50]$	$\%$

Storing

While the BITalino device acquires data, the developed software receives it in batches of 100 samples on the Raspberry Pi, to avoid the computationally expensive processing of one sample at a time. The RPi organizes and concatenates the data from the multiple BITalino devices (if that is the case), resulting in a structure as illustrated in Table 4.2. This data is then saved in the corresponding *.txt* file, which includes a description header that contains all the information needed to interpret the data saved, such as the MAC address of each BITalino, the channels acquired from each, the signal being acquired, and other relevant settings (Appendix A.3 has a more detailed description of the format of the acquisition file).

Table 4.2: Illustration of the structure of the data saved in the acquisition *.txt* file. Example in which the acquisition is performed with 2 BITalino devices: one acquiring channels [A1,A2,A4] and the other [A1,A2]. Note $s_{i(Aj,BITl)}$ corresponds to sample i acquired from channel j of BITalino l.

0	$s_{1(A1,BIT1)}$	$s_{1(A2,BIT1)}$	$s_{1(A4,BIT1)}$	0	$s_{1(A1,BIT2)}$	$s_{1(A2,BIT2)}$
1	$s_{2(A1,BIT1)}$	$s_{2(A2,BIT1)}$	$s_{2(A4,BIT1)}$	1	$s_{2(A1,BIT2)}$	$s_{2(A2,BIT2)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
15	$s_{16(A1,BIT1)}$	$s_{16(A2,BIT1)}$	$s_{16(A4,BIT1)}$	15	$s_{16(A1,BIT2)}$	$s_{16(A2,BIT2)}$
0	$s_{17(A1,BIT1)}$	$s_{17(A2,BIT1)}$	$s_{17(A4,BIT1)}$	0	$s_{17(A1,BIT2)}$	$s_{17(A2,BIT2)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

The first column corresponding to each acquisition device (identified as *"nSeq"*: Sequence Number[5]) is an additional element provided by BITalino with each sample, that can be used to detect issues with the acquisition (e.g. if the sequence is not correct, there are missing samples).

Decimation for Visualization

The first step towards visualization is the decimation of the data, applied to each batch of 100 samples. This procedure is required to allow for proper display of the signals, further explained in Section 4.1.4. The target sampling rate is 100 Hz, which represents a significant subsampling compared to the usual 1000 Hz used for physiological signals, while at the same time preserving the most significant frequencies of each type of signal.

[5]The Sequence Number provided by BITalino, along with the samples, overflows ate 15, hence following a sequence from 0 to 15.

The decimation was performed in Python using the function *decimate()*, from the *scipy.signal* module, using the default parameters (an order 8 Chebyshev type I IIR filter, with phase shift prevention). This technique differentiates from regular subsampling (i.e. removing samples from the signal) by avoiding aliasing. Aliasing occurs when the signal contains frequencies higher than the Nyquist frequency (half the sampling rate) and, therefore, the sampling can not retain all the characteristics of the signal, causing distortion[6]. To do so, the signal is filtered using a low-pass filter that limits the frequency content of the signal to below the Nyquist frequency, and only then it is subsamples.

Plotting

After being decimated, the data is sent to the mobile app, through the MQTT channel, as a list of lists, in which each sub-list contains the samples of a single channel (as illustrated in Figure 4.1, using the same example as in Table 4.2).

$$\begin{aligned}
&[[s_{1(A1,BIT1)}, s_{2(A1,BIT1)}, ..., s_{16(A1,BIT1)}, s_{17(A1,BIT1),...}],\\
&[s_{1(A2,BIT1)}, s_{2(A2,BIT1)}, ..., s_{16(A2,BIT1)}, s_{17(A2,BIT1),...}],\\
&[s_{1(A4,BIT1)}, s_{2(A4,BIT1)}, ..., s_{16(A4,BIT1)}, s_{17(A4,BIT1),...}],\\
&[s_{1(A1,BIT2)}, s_{2(A1,BIT2)}, ..., s_{16(A1,BIT2)}, s_{17(A1,BIT2),...}],\\
&[s_{1(A2,BIT2)}, s_{2(A2,BIT2)}, ..., s_{16(A2,BIT2)}, s_{17(A2,BIT2),...}]]
\end{aligned} \tag{4.1}$$

The data received by the mobile app is then be used to populate the corresponding plot (which is implemented using a Canvas, in which the samples are drawn) in the visualization page. Each plot has a corresponding dataset, which is a list of fixed length, whose content is drawn on the Canvas every time a new sample is added to the dataset (Canvas Widget is rebuilt). Every time a new set of samples is received, each channel (i.e. each sub-list) is gone through and each sample point is individually added to the end of the corresponding dataset.

In order to maintain the length of the datasets, whenever this length is exceeded, the first sample of that dataset is removed before the Canvas gets rebuilt. This value is in fact dependent on the device hosting the mobile app, as it corresponds to the available width for the Canvas (i.e. the number of samples corresponds to the number of pixels). By setting it this way, we guarantee optimal adaptability for every mobile device and, most importantly, that there is not an excess amount of points being drawn without any practical effect on visualization. With this implementation, the width of the Canvas, establishes the time interval of the signal that the user can visualize at any given instant (along with the chosen sampling rate), as described in (4.2). This is why subsampling is a crucial step for visualization; for a Canvas with $width = 320$, if the original sampling rate (1000 Hz) was kept, only 0.302 seconds would be plotted, whereas a sampling rate of 100 Hz would provide us 3.02 seconds of signal.

$$t_{interval} = \frac{width_{(canvas)}}{sampling\ rate} \tag{4.2}$$

[6]Aliasing can also be explained in the frequency domain: when the signal is sampled at an inadequate frequency, there is

Another important matter when plotting the data is the range of each plot; whenever the sensors are discriminated by the user, this information is used to trigger the use of predefined ranges according to the correct physical units (see Table 4.1 for the ranges). In case the type of signal is not provided (and hence the data ranges between 0 and 1023), the range of the plot is set dynamically, making sure that the whole signal being displayed is visible at all times and that the range is never disproportionately large, when compared to the data[7].

Manual Annotations

To facilitate post-processing of the data, the app allows for manual annotations while recording. As already mentioned, EEG recordings are commonly extremely long, hence why they are usually partitioned into 1 hour long files, however this implies that all files are reviewed if one particular event is to be found.

Manual annotations, despite being only individual temporal markers, are extremely valuable to identify which events exist within a file, in a standardized, computationally-readable format. Example use-cases are identification of known artifacts, such as battery changes or equipment misplacement, for easy data cleaning; or seizure events, which is particularly useful in the context of epilepsy, since many files have no seizures whatsoever, having little interest for epilepsy-related algorithms. Moreover, this annotation setup is flexible enough to be implemented in the patient-version of the app, for future applications. Appendix A.3 illustrates the format of the annotations and Appendix A.2 describes how this is achieved within the mobile app.

4.2 Characterization of *EpiBOX*

In order to perform a preliminary evaluation of the proposed system, a testing plan was devised, having two principal components: technical characterization and usability.

4.2.1 Technical Characterization

The technical characterization evaluates some key characteristics of *EpiBOX*, both in the domain of the mobile app and in the acquisition system as whole. Table 4.3 summarizes the estimated characteristics of *EpiBOX* and is divided in three sections: mobile app, acquisition session and communication properties. The values for the first two sections regard a trial acquisition of approximately 10 hours (with $\sim$5 minutes of active screen time[8], with the acquisition of the maximum number of 12 channels. The last section concern inherent communication properties of *EpiBOX* (estimated for optimal conditions, without any physical barriers).

Regarding the acquisition properties analyzed, it is possible to recognize that the acquisition was very

overlap of the copies of the signal spectrum, adding them together, which causes corruption of the frequency content and distorts the signal [60]

[7]ACC, PPG and SpO_2 data follow this same implementation, since they are not transformed to their correct physical units.

[8]Note that the mobile device does not need to have an active screen for the acquisition to progress, only when visualization is necessary, which saves in energy and memory consumption.

Table 4.3: Characteristics obtained for *EpiBOX*, where the first two sections correspond to a trial acquisition of approximately 10 hours (with ∼5 minutes of active screen time) and the last one corresponds to inherent communication properties of *EpiBOX* (estimated for optimal conditions, without any physical barriers). Note: Bluetooth range was not measured, but rather it is given in BITalino documentations, in the official website. Note: frame rendering values correspond to average values computed for some minutes only.

Mobile App	**Memory usage**	App: 87.97 MB Data: 46.55 MB
	Energy consumption	17.93 mAh
	Frame rendering	w/ visualization: 30.3 ms w/o visualization: 9.9 ms
Acquisition	**Acquisition duration**	10 h, 15 min and 35.687238 s
	Total # samples	Expected: 36 935 687 Actual: 36 827 800
	# Lost samples in batch	0
	# Files	Expected: 11 Actual: 12
	File durations	(51 $\pm$ 16) min and (9 $\pm$ 19) s
	File sizes	(110 $\pm$ 41) MB
Commun.	**Bluetooth range**	∼10 m
	WiFi speed	72 Mbps
	WiFi strength	-30 dBm

efficient, particularly considering the large number channels, losing only 107 887 samples (0.3%) during processing times (e.g. opening and closing files, saving data and eventual mid-session re-connections to the acquisition devices), which corresponds to less than 2 minutes of non-covered acquisition time. Furthermore, it is relevant to notice that no samples were lost during acquisition or during the data transfer from the BITalino devices to the RPi (given by number lost samples in batch). Additionally, this acquisition allows to illustrate the behavior of *EpiBOX* whenever an unexpected event takes place (e.g. disconnection of one of the devices), as it closes the current acquisition file and opens another one, once the acquisition is resumed, allowing for easier identification of acquisition issues (which can be verified by comparing the number of expected number of files - one per hour - and the actual number).

Overall, *EpiBOX* exhibits satisfactory properties, particularly regarding the communication channels, adequate for application in a hospital/home environment. One of the advantages of *EpiBOX* that is not addressed in Table 4.3 is the security of the data transfer, considering it is executed through a private network, not connected to the internet. Nonetheless, there is one property that can still be improved: to achieve an adequate rendering rate of 60 Frames per Second (FPS), each frame must have a maximum rendering time of 16 ms [9], which is not the case in frame rendering with visualization. This results in UI jank (i.e. not smooth visualization, with skipped frames), which does not have a huge impact for the purpose of the visualization, but should be improved for user experience.

Another interesting characteristic to analyze is the time it takes to share a message between the two devices: assuming a WiFi speed of 72 Mbps, we can compute this through the size of the MQTT publish

[9] Flutter. "Using the Timeline View". [Online].

packet. As was described in Section 2.3.2, an MQTT message packet is composed by a fixed header with 2 bytes; a variable header which, for publishing a message, corresponds to 2 bytes as well; and a payload of variable length, corresponding to the actual message. Regarding the transfer of the acquired data for visualization, this corresponds to a maximum total packet size of 743 bytes, corresponding to the acquisition of 12 channels, with 10 samples each (Appendix A.4 provides the packet sizes for all the remaining messages exchanged through MQTT). Considering the publishing and download of this message, the total time needed for the transfer amounts to $\frac{(743 \times 2)\ B}{72\ Mb} = 0.165\ ms$[10], which surpasses the frequency of data transfer from the BITalino to the RPi.

4.2.2 Usability

Ideally, *EpiBOX* should be directly compared to its predecessor, i.e. the system that was previously used as a long-term biosignal acquisition system in the scope of the PhD research already mentioned in the Introduction. However, only one person could perform this comparison, since the system required handling by someone proficient, hence, such a comparison analysis would not be considered statistically relevant. Instead, in order to assess the usability of *EpiBOX*, a preliminary testing phase will be carried out with the collaboration of the units of neurology, at Hospital de Santa Maria and Hospital Egas Moniz. This experiment has two main goals:

1. Assess the usability of *EpiBOX* by medical staff, without any assumption on their proficiency in the use of modern technology; and consequently evaluate the potential of *EpiBOX* to be integrated in a normal clinical context, as a collaboration tool between researchers and clinical staff.

2. Identify potential improvements and adaptations needed for an optimal operation and usability of the system.

To achieve this, a post-session questionnaire was designed with two independent sections, each conceived to answer one of the goals. The first section is based on the widely-used *System Usability Scale*, which provides a reliable, industry-standard tool to evaluate the ease of use of a large variety of products and services. The second section aims at identifying potential technical problems on a real-world application environment (instead of a development environment), as well as system design flaws that provide a non-optimal experience for the user and should therefore be revised. The Portuguese version of the questionnaire can be consulted in Appendix B.

However, due to COVID-19 restrictions, the testing phase is only expected to start in the beginning of January, hence why it will not be possible to analyze the preliminary results within the context of this dissertation.

[10] B: byte, Mb: megabit, 1 Mb = 125000 B.

4.3 EEG Acquisition Wearable

Part of the purpose of this work is to develop a seizure detection algorithm that can be integrated with *EpiBOX* for a continuous monitoring and simultaneous event detection. As was addressed in Section 2.2.1, EEG is a gold-standard in the analysis of epilepsy, particularly in the detection and prediction of seizures. As such, we required a wearable device for the acquisition of this biosignal, that could be directly implemented with *EpiBOX*.

EmotAI's headband was specifically designed for Esports analytics (namely cognitive, emotional and behavioral), however, it has potential for use in other areas, such as epilepsy. It provides the recording of 2-channel EEG, through 2 pairs of dry electrodes positioned to acquire the channels corresponding to *Fp1* and *Fp2*. It also has an additional electrode, connected to the left ear, that acts as reference and allows to decrease noise and artifacts (the patterns that are common between the reference point and the main electrodes), as well as a Photoplethysmography (PPG). See Appendix A.1 for a picture of the headband, along with the positioning of the electrodes and interior of the device.

The limited-channel configuration that EmotAI provides is particularly suitable for the goals of this work as it enables a relatively discreet approach for continuous monitoring of EEG. Moreover, EmotAI's headband is based on BITalino, which guarantees the necessary scientific rigor of measurements, as well as a seamless integration with *EpiBOX* without any adaptation needed.

Throughout the next Chapter, all design decisions were made having in mind the implementation of the algorithm using EmotAI's headband as the acquisition device, namely regarding the choice of channels and re-referencing of the dataset, as described throughout the coming sections.

Chapter 5

Seizure Detection Algorithm

Due to COVID-19 restrictions, it was not possible to collect real-world data in the context of this work, hence, an external dataset was used (and adapted) to act as a proof of concept of the acquisition of EEG data using EmotAI's headband, with simultaneous automatic seizure detection.

5.1 Development

5.1.1 Data and Preprocessing

Figure 5.1 shows a schematic of the preprocessing steps that will be described hereafter.

Figure 5.1: Schematic of the preprocessing steps performed for each of the recordings in the EDF files in TUSZ dataset.

Dataset

The TUH EEG Seizure Corpus (TUSZ) [4] was used, one of the largest publicly available archive of clinical scalp-EEG, created by Temple University Hospital. The use of this database was only possible due to the participation in the Neureka 2020 Epilepsy Challenge. This corpus was specifically curated towards the training (and testing) of seizure detection algorithms, by selecting the most relevant segments of data from the much larger TUH EEG Corpus [61]. The database consists of approximately 504 hours of scalp-EEG recordings ($\sim 7\%$ corresponding to seizures), from 692 patients and containing more than 3000 seizures, all manually annotated.

The recordings in this corpus were performed according to the widely used 10-20 system for electrode placement, using 19 electrodes, with two separate unipolar montages: Average Reference (AR) and Linked-Ears Reference (LE). The AR montage uses the average of a certain number of electrodes as the reference, whereas LE uses a lead adapter to link both ears, using this as reference [62]. The

corpus provides three separate datasets: *01_tcp_ar*, *02_tcp_le* and *03_tcp_ar_a*, which are described in Table 5.1. Table 5.2 presents the 19 channels reported for montages *01_tcp_ar* and *02_tcp_le* (montage *03_tcp_ar_a* has the same channels except for the auricular ones).

Table 5.1: Datasets provided by TUSZ v1.5.2. The number of channels reported concern the channels labeled in Table 5.2.

Dataset	Montage Format	# Channels
01_tcp_ar	AR	19
02_tcp_le	LE	19
03_tcp_ar_a	AR	17

Table 5.2: Channel labels from the TUSZ v1.5.2 dataset. Each channel corresponds to the difference in voltage between the electrode from the 10-20 system montage and the reference (REF) of each dataset, i.e. AR or LE.

Frontal	F3-REF	F4-REF	F7-REF	F8-REF
Pre-Frontal	Fp1-REF	Fp2-REF		
Parietal	P3-REF	P4-REF		
Temporal	T3-REF	T4-REF	T5-REF	T6-REF
Central	C3-REF	C4-REF	CZ-REF	
Occipital	O1-REF	O2-REF		
Auricular	A1-REF	A2-REF		

The annotations (provided as a *.txt* file for each acquisition file) in TUSZ contain 12 different labels, corresponding to background (non-ictal) and seizure (ictal) events. The labels used to annotate seizures are: Focal Non-Specific Seizures (FNSZ), Generalized Seizures (GNSZ), Simple Partial Seizures (SPSZ), Complex Partial Seizures (CPSZ), Absence Seizures (ABSZ), Tonic Seizures (TNSZ), Clonic Seizures (CNSZ), Tonic-Clonic Seizures (TCSZ), Atonic Seizures (ATSZ), Myoclonic Seizures (MYSZ), and Non-Epileptic Seizures (NESZ). Non-seizure events are annotated as Background (BCKG). Table 5.3 summarizes the number of instances reported for each seizure type and for each dataset.

Table 5.3: Summary of the number of events annotated for each dataset. Note: originally, TUSZ was split in a training dataset and a validation dataset, however these were combined so that the split could be optimized for this work. FNSZ: Focal Non-Specific Seizures, GNSZ: Generalized Seizures, SPSZ: Simple Partial Seizures, CPSZ: Complex Partial Seizures, ABSZ: Absence Seizures, TNSZ: Tonic Seizures, CNSZ: Clonic Seizures, TCSZ: Tonic-Clonic Seizures, ATSZ: Atonic Seizures, MYSZ: Myoclonic Seizures, NESZ: Non-Epileptic Seizures.

	01_tcp_ar	***02_tcp_le***	***03_tcp_ar_a***
FNSZ	1070	231	538
GNSZ	428	87	68
SPSZ	52	-	-
CPSZ	138	83	146
ABSZ	2	97	-
TNSZ	62	-	-
CNSZ	-	-	-
TCSZ	28	16	4
ATSZ	-	-	-
MYSZ	1	2	-
NESZ	-	-	-

Channel Extraction and Re-referencing

As mentioned, the algorithm will be designed as a proof of concept of seizure detection using EmotAI's headband, therefore only two channels will be extracted from the acquisition files, stored in European Data Format (EDF): *Fp1-REF* and *Fp2-REF*. Since both datasets *01* and *03* share the same reference point and both include the channels of interest, these two will be combined in a single dataset, hereinafter referred to as *0103*.

Considering that the reference point is not the same in the dataset and the headband, offline re-referencing should be performed in order to guarantee that each channel represents the same relationship in both cases (it is important to remember that, ultimately, the signal from a channel is the difference in voltage between two electrodes, in two regions of the scalp [63]).

In theory, this re-referencing to a common montage would be enough to cancel the effects of the different reference points; however, it is important to note that, in practice, this might not be the case: the reference point can have a major impact on the nature of the waveforms, due to the nonlinearity of the brain and scalp conduction paths [63]. As such, the potential effect of the initial reference will be analyzed during this work. Moreover, we must take in consideration that this can have unpredictable effects once we wish to transpose the algorithm into the headband, which will be a target of future work.

For all acquisition files, in both datasets (*0103* and *02*), the channels *Fp1-REF* and *Fp2-REF* were extracted and re-referenced to *Fp1-Fp2*, as in (5.1), timestamp-wise, resulting in a single channel.

$$(Fp1 - REF) - (Fp2 - REF) = Fp1 - Fp2 \tag{5.1}$$

Denoising

It is known that filtering can cause significant distortion of the signal, hence why many researchers choose to skip this preprocessing step [64]. However, low-frequency noise is usually the predominant source of noise in electrophysiological data [64] and can have a major impact in the extraction of features (particularly spectral ones). In fact, in Figure 5.2(a) (which shows the Power Spectral Density (PSD) of one of the recordings), one can see that the spectral power is highly concentrated in the range of very low frequencies, diminishing all other frequencies. Therefore, a highpass filter is almost crucial to increase Signal-to-Noise Ratio (SNR) and allow a proper spectral analysis of the signals.

All recordings were filtered using an 8-order highpass filter, with cutoff frequency of 0.8 Hz and an 16-order lowpass filter, with cutoff frequency of 48 Hz. Note that the choice of 0.8 Hz for the cutoff frequency is slightly above the recommended in [64]. However, it was a conscious decision, as it allows to not only remove the unwanted, low-frequency, noise components, but also to approximate the range of frequencies of the TUSZ recordings to the one EmotAI's headband is able to acquire (considering its bandwidth)[1].

[1] BITalino's EEG sensors are limited to the bandwidth of 0.8-48 Hz, as described in the sensor's official datasheet (on BITalino's website).

Figure 5.2(b) shows the PSD of the same recording after being filtered. From the significant decrease in maximum power between the two PSD, one can ascertain an effective reduction of low-frequency noise. Moreover, the power-line interference at 60 Hz is also successfully removed by the filter technique applied.

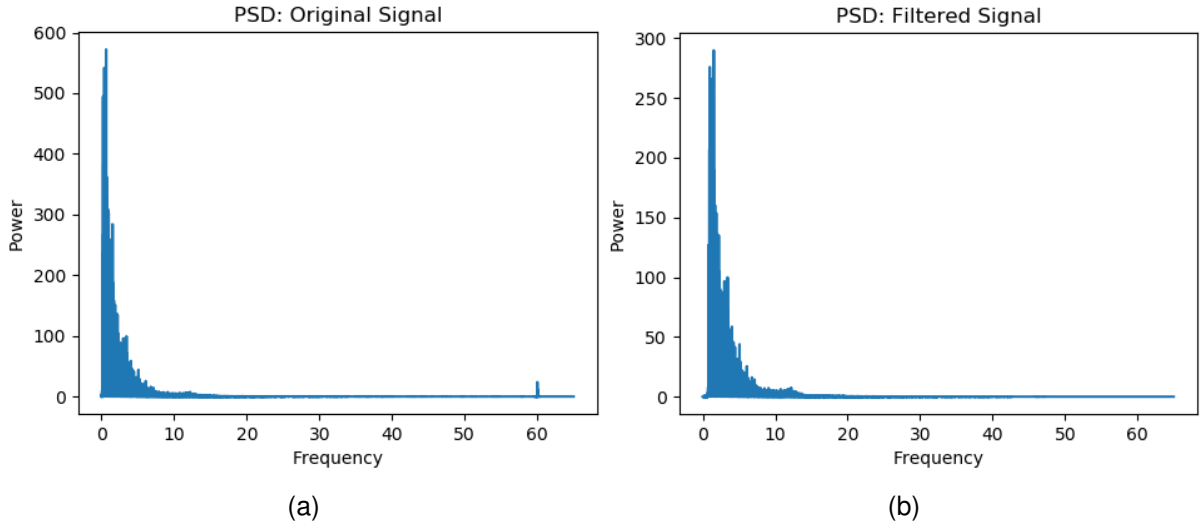


Figure 5.2: PSD of the same recording (a) before any filtering technique was applied and (b) after being filtered with a highpass and a lowpass filter, at cutoff frequencies of 0.8 Hz and 48 Hz, respectively. Note: only showing a portion of the original range of 0-125Hz, since the remaining frequencies are not considered of interest for this work and had negligible powers.

Regarding artifacts, within the annotation procedure of the TUSZ dataset, besides the labels used for seizure-related events and background, there were also available six other possible labels, used to identify specific artifact annotations [12]: eye movement (EYEM), chewing (CHEW), shivering (SHIV), muscle artifact (MUSC), electrode pop (ELPP), and electrostatic artifact (ELST). However, within the annotation files provided, none of these labels were found, which suggests that only previously cleaned data was included. Therefore, artifact detection and removal will not be addressed in the scope of this work.

Signal Segmentation

Each recording was split into equally-sized epochs, from which a feature vector would then be computed. Regarding the length of the epochs, there is some degree of variety across literature, where it is suggested a trade-off between avoiding non-stationarity (too long epochs) and not overlooking important information (too short epochs that can not capture lower frequency patterns). However most agree on a reasonably short epoch. For example, neurologists typically analyze EEGs in windows of 10 seconds and identify events with a temporal resolution of $\sim$ 1 second [46]. Moreover, aiming at a real-time detection algorithm, the length of the epochs evidently influences the time resolution of the alerting mechanism, which is another reasoning for the choice of the parameter.

There is some literature regarding the optimal window size for signal segmentation in seizure detection, however, to the author's best knowledge, there are no studies focusing on the specific channel

configuration used in this work, which could have a major influence in the performance. Therefore, it would be interesting to take this factor into consideration within the design of the algorithm, i.e. testing different epoch lengths (e.g. [0.25, 0.5, 1, 2, 3, 5] seconds), with the purpose of determining the optimal length for this particular channel configuration. However, the feature extraction process proved to be too time-consuming, due to the size of the dataset and the computation time for some of the features. Hence, instead of performing this optimization study, each recording will only be split into non-overlapping epochs with duration of 1 second[2].

5.1.2 Feature Extraction

As was demonstrated in Section 3.2.1, there is great variability in the features used for the task of seizure detection across literature. Besides, since the state-of-the-art lacks in-depth studies regarding the most informative features specifically towards limited-channel detection, a more comprehensive set of features will be extracted.

Seizure detection is a typical case of the curse of dimensionality described in Section 2.2.2. If the features are not evaluated properly, due to the class imbalance and low number of training points, we may be risking biasing the classifiers with information that does not contribute to the detection of seizures. In order to prevent this, in Section 5.1.3, a feature selection process will allow to eliminate the features that are less informative and/or contribute less for the task of seizure identification.

The complete set of extracted features for each epoch (enumerated in Table 5.4) corresponds to a single feature vector concatenated to form a single array of dimension $1 \times \#features$. Note that the number of relative energies and statistics of the coefficients varies according to $dlvl$ (decomposition level), which in this case corresponds to 5 (as explained below).

Table 5.4: List of features extracted for each epoch in the scope of this work. Note: for the DWT-based features, their computation is explained in Section 5.1.2. FD: Fractal Dimension, SampEn: Sample Entropy, HE: Hurst Exponent, Std: Standard deviation. cA*n*: approximate coefficients from level *n*, cD*n*: detail coefficients from level *n*.

Non-linear	**DWT-based**				
	Relative energies	**Mean coeff**	**Std coeff**	**Kurtosis coeff**	**Skewness coeff**
SampEn	re_{cD2}	mean_{cAdlvl}	std_{cAdlvl}	kurt_{cAdlvl}	skew_{cAdlvl}
Higuchi FD	$\vdots$	mean_{cDdlvl}	std_{cDdlvl}	kurt_{cDdlvl}	skew_{cDdlvl}
HE	re_{cDdlvl}	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	re_{cAdlvl}	mean_{cD2}	std_{cD2}	kurt_{cD2}	skew_{cD2}

Non-Linear Features

Considering the nonlinearity of EEG, it is important to include features that are able to capture this innate behavior of biosignals. All of the three nonlinear features used in this work are computed directly from time-series, i.e. from the time-domain of EEG signals.

Entropy-related features measure the unpredictability of a time series, with larger values of entropy

[2]Nevertheless, this analysis would be relevant to perform in future work. Perhaps, the present analysis can be used to define a smaller set of features to extract, thus decreasing the time of feature extraction and, consequently, enabling this future study.

unveiling less order and more randomness in the signal [65]. Sample Entropy (SampEn), in particular, is less sensitive to noise than other entropy measures and is suitable for short-length time series [66], which is crucial, given the epoch length considered in this study. This measure is obtained with a rather complex algorithm that can be computed as in [67].

Higuchi's Fractal Dimension is a method for computing Fractal Dimension (FD), a nonlinear measure of signal complexity. Complexity is related to synchrony, with increased synchronous oscillations (such as in seizure events) translating to lower values of FD [68]. This measure can be obtained as described in [69].

Lastly, Hurst Exponent (HE) is a measure of long-term memory of the time-series, in the sense that it evaluates the persistence of a trend or pattern across time [70]. The estimation of HE can be computed as in [70].

DWT-Based Features

For the extraction of DWT-based features, each epoch was decomposed using a Daubechies 4 (db4) mother wavelet, which is one of the most commonly used in literature (e.g. [25, 27, 71]). In order to extract sub-bands close to the renowned Delta ($0.1 \leq f < 4\ Hz$), Theta ($4 \leq f < 8\ Hz$), Alpha ($8 \leq f \leq 12\ Hz$), Beta ($12 \leq f \leq 30\ Hz$) and Gamma ($30 \leq f \leq 70\ Hz$), the decomposition was performed until level 5, as illustrated in Figure 5.3 (considering a sampling frequency of 250 Hz). Note, however, that this might not be possible if smaller epoch lengths were considered: according to the number of samples in the epoch and the length of the mother wavelet, there is a maximum level of decomposition (given by (5.2)), above which all coefficients become corrupted by edge effects, caused by signal extension[3].

$$max\ level = \log_2(\frac{length(epoch)}{length(filter-1)}) \tag{5.2}$$

Table 5.5 shows the maximum level of decomposition for a set of epoch durations and what would be the resulting (total) number of features.

Table 5.5: Maximum level of wavelet-based decomposition, for a db4 wavelet (of length 8), considering the specified epoch duration and a sampling frequency of 250Hz. Total number of features for each epoch duration according to the decomposition level.

Epoch duration [s]	Maximum level	Decomposition level ($dlvl$)	Nb. features
0.25	3	3	18
0.50	4	4	23
1.00	5	5	28
2.00	6	5	28
3.00	6	5	28
5.00	7	5	28

The relative sub-band energies (re$_{cDn}$ and re$_{cAn}$) were computed, based on (5.3), for each of the frequency ranges of interest, i.e. the ranges corresponding to $cD2$-$cDdlvl$ and $cAdlvl$. To do so, the

[3]PyWavelets. "Discrete Wavelet Transform (DWT)". [Online].

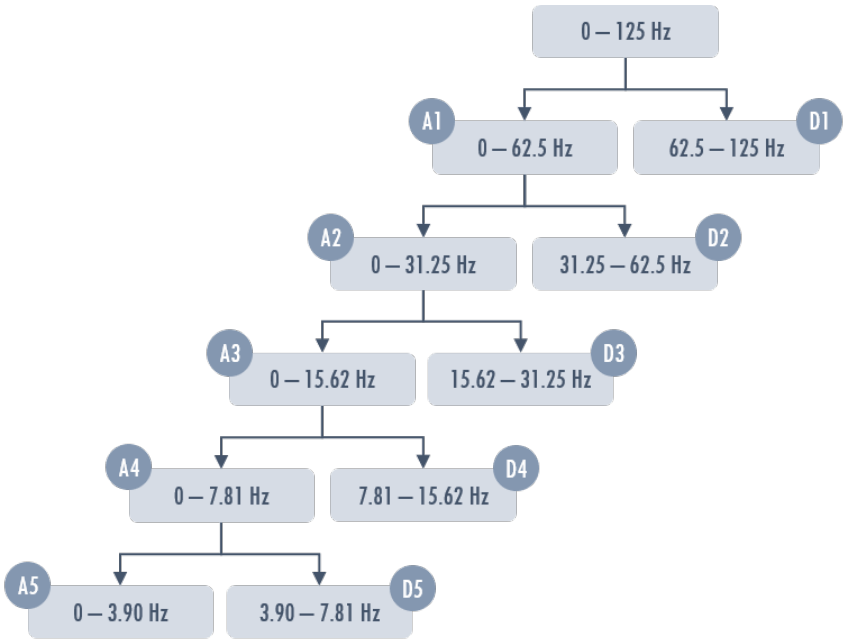


Figure 5.3: DWT multilevel decomposition, considering a sampling frequency of 250 Hz, according to the decomposition levels and resulting frequency ranges. A*n*: approximation level *n*, D*n*: detail level *n*.

signal was reconstructed using each set of coefficients individually, resulting in a reconstruction of the epoch's content corresponding to that bandwidth. Therefore, when the energy of the signal is computed, it corresponds to the energy of that frequency sub-band.

$$E[sub\ band] = (x_{cD/An})^T . x_{cD/An} \tag{5.3}$$

The remaining features were extracted directly from the coefficients, i.e. statistical measures (mean, standard deviation, kurtosis and skewness) of the distribution of each set of coefficients, computed as shown in (5.4), (5.5), (5.6) and (5.7). Knowing that the coefficients correspond to the correlation of the signal at a specific time frame with a scaled version of the mother wavelet, analyzing their distribution can provide valuable information on the overall behavior of the signal across that frequency sub-band.

$$Mean,\ \mu\ (measure\ of\ central\ tendency) = \frac{1}{N}\sum_{k=0}^{N-1} c_k \tag{5.4}$$

$$Standard\ deviation,\ \sigma\ (measure\ of\ spread) = \sqrt{\frac{1}{N-1}\sum_{k=0}^{N-1}(c_k-\mu)^2} \tag{5.5}$$

$$Kurtosis\ (measure\ of\ "tailedness"^4) = N\frac{\sum_{k=0}^{N-1}(c_k-\mu)^4}{(\sum_{k=0}^{N-1}(c_k-\mu)^2)^2} - 3 \tag{5.6}$$

$$Skewness\ (measure\ of\ asymmetry) = \frac{\frac{1}{n}\sum_{k=0}^{N-1}(c_k-\mu)^3}{(\frac{1}{n}\sum_{k=0}^{N-1}(c_k-\mu)^2)^3} \tag{5.7}$$

5.1.3 Sample Preprocessing and Feature Selection

Figure 5.4 shows a schematic of the steps performed for sample preprocessing and feature selection, applied to the samples acquired during feature extraction, which will be described hereafter.

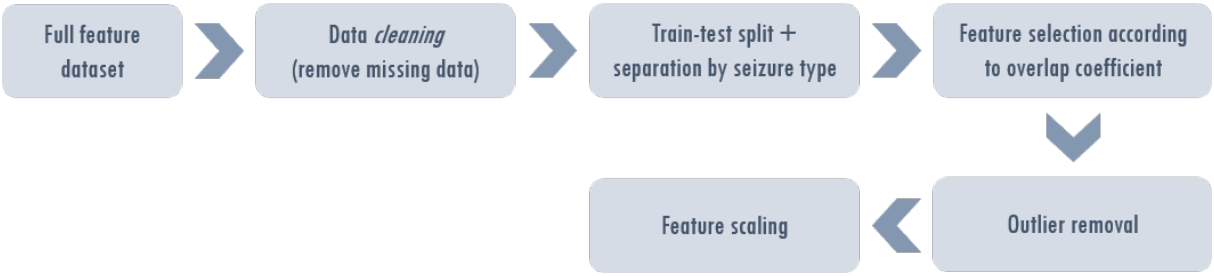


Figure 5.4: Schematic of the steps performed for sample preprocessing and feature selection, applied to the samples acquired during feature extraction.

Train-Test Split

Before any analysis was performed, the data was "cleaned", by removing all samples that included any missing data (i.e. `NaN` values), resultant from incorrect sensor readings or unexpected behaviors during the feature extraction process. This ensures that the ML algorithms will not run into unexpected problems when trying to fit the data.

Train-test splits serve the purpose of dividing the original dataset into two independent subsets that can be used to train a ML algorithm and then test it on unseen data, guaranteeing unbiased results. Targeting the implementation of an online algorithm (that may have temporal constraints in its final classification state), it is preferable to test the performance in an online format as well. However, in order to do so, the test samples must be consecutive and have a temporal dependency. This is not possible if the samples are shuffled pre-splitting. On the other hand, Python *sklearn* `train_test_split` allows for non-shuffled splitting, but this implies that no measures are taken to preserve the proportion of class samples in the new subsets[5]. Hence, an alternative splitting mechanism is proposed, based on an exhaustive search technique:

1. The target number of samples is computed as 20% of the total number of samples belonging to a specific seizure type in the original dataset (note that the train-test split was performed separately for each seizure type).

2. For each recording session, the number of samples belonging to background and a specific seizure type is calculated.

3. All the possible combinations of 1 or 2 acquisition sessions are computed, along with the total number of resulting samples for both classes.

4. The optimal combination of sessions is selected as the one that has the smallest difference to the target number of samples for both classes, as given by (5.8)[6].

[5]Train-test splits in which the proportion of class samples is preserved are known as *stratified splits*. If not applied, there is no guarantee that there will be a representative number of samples of all classes in both subsets.

[6]In some cases, the optimal combination was not satisfactory (with a ratio of seizure to background samples smaller that 2:3),

$$\underset{s_1,s_2}{\arg\min}\,[N_{target}-(N_{bckg}s_1+N_{bckg}s_2)]+[N_{target}-(N_{sz}s_1+N_{sz}s_2)]+[N_{train\_target}-(N_{train\_sz}s_1+N_{train\_sz}s_2)] \tag{5.8}$$

where

N_{target}: target number of samples for both classes (in the test subset)

$N_{train\_target}s_i$: target number of samples for seizure class (in the train subset)

$N_{bckg}s_i$: number of samples belonging to background in session i (in the test subset)

$N_{sz}s_i$: number of samples belonging to a specific seizure type in session i (in the test subset)

$N_{train\_sz}s_i$: number of samples belonging to a specific seizure type in session i (in the train subset)

This method yields two independent subsets for each seizure type, containing background and seizure samples, whose patterns are unique between train and test subsets. In fact, with this splitting technique, each acquisition session is used exclusively in one of the subsets. Table 5.6 exemplifies the result of this method for both datasets. Note that this technique was designed to ensure a balanced test set (i.e. approximately equal number of samples belonging to background and seizure classes), with a representative number of seizure samples, instead of aiming at the usual 80%-20% split of the original dataset, so that the whole test subset can be used without further balancing.

Analyzing Table 5.6, it is possible to conclude that the method applied to split the original dataset was, overall, satisfactory, achieving balanced subsets and close to the target number of samples. However, in some cases no combination of sessions or files were able to achieve the desired split (e.g. FNSZ from *0103*), whose splits were either unbalanced or did not respect the target number of samples. Moreover, ABSZ and MYSZ from *0103* did not have enough seizure samples, hence why they were not included in the analysis.

Table 5.6: Summary of the number of test samples after the splitting procedure described above. The first row indicates the dataset to which the subsets belong. With these values we can infer how balanced the test subsets are and how close they correspond to 20% of the seizure samples.

	0103			**02**		
	bckg	**sz**	**target**	**bckg**	**sz**	**target**
FNSZ	4149	5355	20849	1830	2894	3206
GNSZ	5321	5301	8820	2952	5290	2953
SPSZ	476	248	424	–	–	–
CPSZ	5662	4237	5777	1367	1269	1245
ABSZ	–	–	3	1089	139	159
TNSZ	229	69	235	–	–	—
TCSZ	517	436	448	723	542	658
MYSZ	–	–	4	628	606	258

in which cases, the same method of optimal combination was applied to the files that contained that seizure type (instead of sessions), in order to have more flexibility over the combinations.

Feature Selection

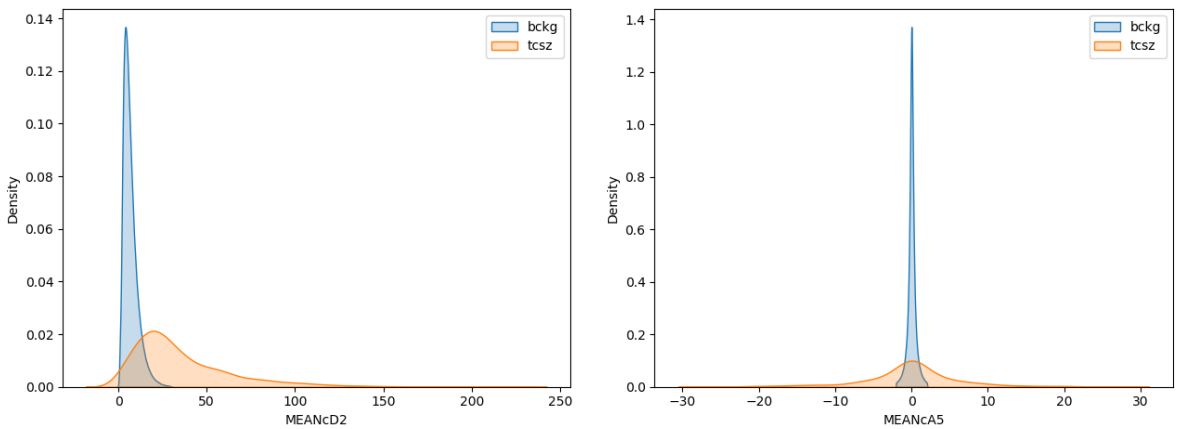
Note that all the procedures done herein after, within the Section of 5.1.3, were performed only on the *train subset*, separately for each *seizure type*.

Feature selection allows to reduce the number of input variables to consider in the ML model. At this stage, it is important to identify the features with the largest predictive power *a priori* (i.e. without performing any classification) because, otherwise, the feature selection would have to be performed entirely with more computationally expensive methods, such as wrapper methods [72]. It is important to note, however, that the "best" individual features do not always comprise the optimal set of features, so this process can be somewhat incomplete and misleading[7]. For the purpose of this work, two main criteria were used: linear separability and differential spread.

Linear separability: when there is a large difference in the distributions' means.

Differential spread: when there is a large overlap between the sets, but one of the classes is highly concentrated (i.e. small variance) and the other is highly spread out (i.e. large variance).

These two criteria can be inferred through visual analysis of the simultaneous plot of the univariate distributions corresponding to both background and seizure samples (separately for each feature). Figure 5.5 shows a Kernel Density Estimate (KDE)[8] plot, in which we can see two good examples of (a) linear separability, considering the distributions have different means and only a small portion overlaps, and of (b) differential spread, in which we can see that both distributions share the same mean, but have very distinct spreads.



(a) (b)

Figure 5.5: KDE plot of the univariate distributions corresponding to both background and TCSZ for features (a) mean$_{cD2}$ and (b) mean$_{cA5}$. Extracted from dataset *0103*.

However, in order to have an objective measure of these criteria, an *overlap coefficient* was computed. Statistical methods that assume specific conditions (e.g. normal distribution) are not suitable for this particular application, given that it can not be guaranteed for all the variables under study. Hence, an alternative version of an overlap coefficient is proposed:

[7]However incomplete the proposed filter-based selection method is, it is the most practical at this stage.

[8]Kernel Density Estimate is a method for visualization of univariate (or bivariate) distributions, which represents an estimate of the distribution of a variable, based on a kernel smoothing of the histogram representation of the data.

1. For each variable (i.e. feature), two normalized histograms are computed, one corresponding to the background class and the other to the seizure class, in which the *range* and number of *bins* are the same for both estimates.

2. The histograms of both distributions are multiplied bin-wise and the total sum is computed, corresponding to the overlap coefficient, as shown in (5.9).

3. The larger the overlap between the two distributions, the larger the coefficient, i.e. if the distributions were equal $c_{overlap} = 1$, and if they did not intersect $c_{overlap} = 0$.

$$c_{overlap} = \sum_{k=0}^{N} \sqrt{(height(B_k^{bckg}) \times width(B_k^{bckg})) \times (height(B_k^{sz}) \times width(B_k^{sz}))} \tag{5.9}$$

where

N: total number of bins in the histogram

B_k^{bckg}: bin *k* from the histogram corresponding to the background samples

B_k^{sz}: bin *k* from the histogram corresponding to the seizure samples

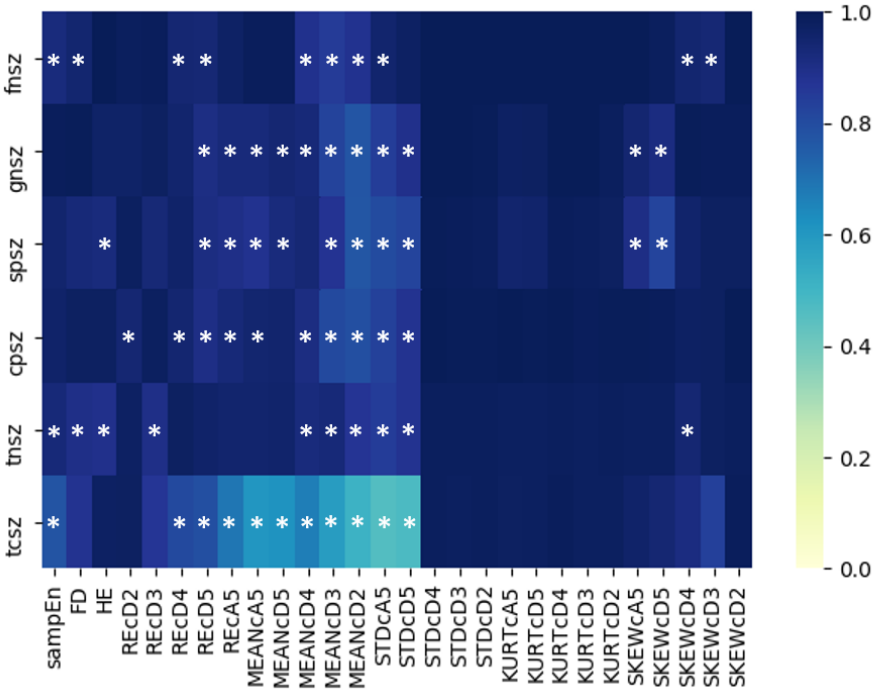
$(height(B_k) \times width(B_k)$: area of bin *k*

The results obtained with dataset *0103* and *02* are displayed in Figure 5.6 (a) and 5.6 (b), respectively. As it is possible to observe, there is no obvious agreement between the features with larger apparent predictive power, within the different feature types. In fact, while some features seem to show some degree of consistency across seizure types, such as mean$_{cD3}$, mean$_{cD2}$ and std$_{cA5}$, overall there is great variability. Even within the two datasets, we can identify some contrasting results, particularly regarding skew$_{cA5}$, skew$_{cD5}$ and skew$_{cD4}$.

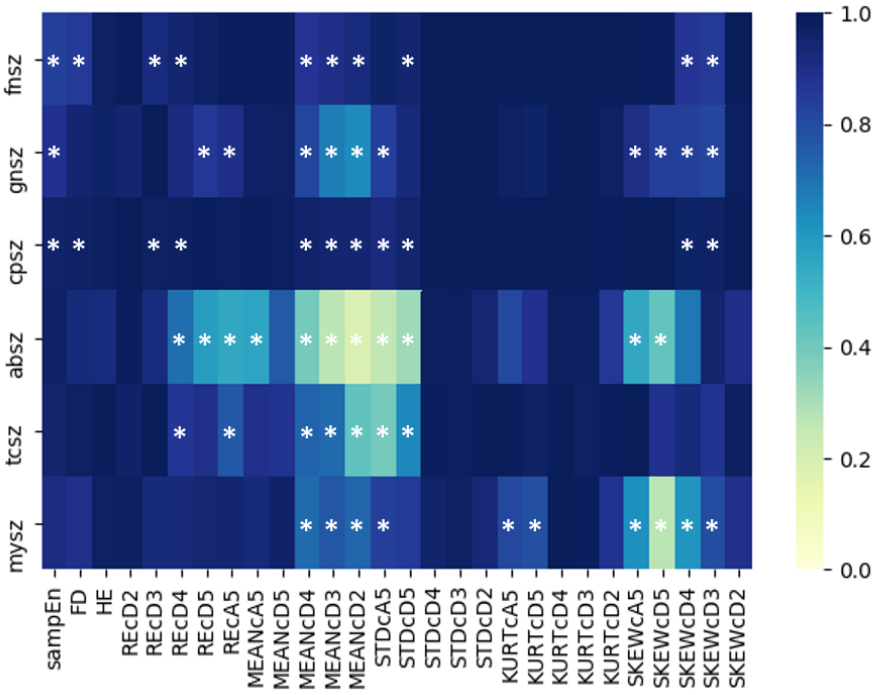
CPSZ is the seizure type which shows the largest overlap between the two distributions across all features, which is not very promising in terms of their usability for the upcoming task. In fact, no features exhibited a $c_{overlap}$ smaller than 0.9 in the *02* dataset. ABSZ, TCSZ (except between std$_{cD4}$ and skew$_{cA5}$) and MYSZ on the other hand, show much more promising results, as we can see from the consistently lower $c_{overlap}$ values.

Considering this variability across seizure types, it would be too strict to apply the same threshold of inclusion to all seizure types, hence why this process will be individually adapted to each. The cutoff threshold was defined as the average value of the coefficients, which can be consulted in Table 5.7, along with the resulting number of features. The actual set of selected features is identified in the heatmaps of Figures 5.6 (a) and (b), as the ones with a white asterisk.

Interestingly, there is an apparent overlap between selected features between montages and within each montage, although it is more evident in the later. For example, the group of five features between mean$_{cD4}$ and std$_{cD5}$ is practically coincident across seizure types and montages. In future work, this evidence can be used as background knowledge to study the possibility of homogenizing the features extracted for each seizure type, and creating a single detection model for the detection of multiple types of seizure.



(a)



(b)

Figure 5.6: Heatmap with the coefficients of overlap between the distributions of background and the respective seizure. Rows correspond to the different seizure types and columns correspond to set of extracted features. Extracted from datasets (a) *0103* and (b) *02*. The selected features are identified with a white asterisk. Note: the seizure types are not all the same within the 2 heatmaps, please verify the row names.

Outlier Removal

As was explained in Section 2.2.2, outliers can be particularly harmful when scaling with a bounded range is performed prior to training, hence why it is a crucial step in the context of this work.

Table 5.7: Cutoff threshold for the overlap coefficients, specific for each dataset and seizure type. These values define the threshold above which the features are rejected during the feature selection process. *# features* columns show the number of features that were selected after applying the corresponding threshold.

	0103		**02**	
	threshold	**# features**	**threshold**	**# features**
FNSZ	0.97	10	0.95	10
GNSZ	0.95	11	0.91	11
SPSZ	0.93	11	–	–
CPSZ	0.95	10	0.98	11
ABSZ	–	–	0.72	12
TNSZ	0.95	10	–	–
TCSZ	0.83	11	0.88	7
MYSZ	–	–	0.85	10

This procedure was performed using Interquartile Range (IQR), which is a quartile-based measure of data variability. Quartiles divide the rank-ordered set of samples into four equal parts (i.e. each part contains 25% of the samples), with three cutoff points - Q1, Q2 and Q3[9]. This outlier removal technique uses these concepts, as shown in (5.10), to compute the upper and lower threshold cutting values for outlier identification[10]. This was performed separately for background and seizure samples, in order to maintain the uniqueness of both classes. Every sample that contained at least one outlier was removed.

$$\begin{aligned} &IQR = Q3 - Q1 \\ &outliers : x < Q1 - 3 \times IQR \,||\, x > Q3 + 3 \times IQR \end{aligned} \tag{5.10}$$

Feature Scaling

Each feature was standardized (i.e. 0 mean and unit variance) prior to model training, using *sklearn*'s `StandardScaler()`. The train subset was standardized first and the standardization parameters were obtained, in order to apply the exact same procedure to the test subset.

5.1.4 Classifier

Offline Training

After the procedure of splitting the original dataset into train and test subsets, the resulting train subsets were, in most cases, highly unbalanced (with background as the majority class). To manage this issue, we could use some kind of penalization parameter to decrease misclassifications of the minority class. However, with such a large number of samples, training with cross validation would become extremely time consuming and computationally expensive. Hence, the majority class will be downsampled by random sampling.

[9] Stat Trek. "Interquartile Range". [Online].

[10] In (5.10), there is a scaling factor of 3, which determines how much of the data is considered an outlier. Usually, this factor is set to 1.5 (as explained in Towards Data Science. "Why 1.5 in IQR Method of Outlier Detection?". [Online].), however, within this dataset, it was excluding an undesirable amount of samples, hence why it was set to 3 instead.

Support Vector Machines (SVM) are a set of supervised ML methods which have become extremely popular due to their versatility (through the use of different kernels, in Kernel SVM) and effectiveness in high dimensional spaces [11]. Particularly in the field of epileptic seizure detection, SVM have been associated to very satisfactory results, often superior to the state-of-the-art (both in limited-channel and high-density configurations, e.g. [43, 71, 73]).

In the scope of this work, *Scikit-learn* implementation was used, which requires the definition of some hyperparameters[12]:

- **Regularization parameter (C)**: common to all SVM implementations and is used to avoid overfitting, by penalizing highly complex decision surfaces; increasing C corresponds to less regularization.
- **Kernel coefficient ($gamma$)**: defined for polynomial, RBF and sigmoid kernels and establishes how much influence a single training sample has; increasing $gamma$ corresponds to small-range influence.
- **Kernel type**: from polynomial, radial basis function and sigmoid.
- **Degree**: only for the polynomial kernel.

For the purpose of this work, the Radial Basis Function (RBF) kernel will be used (as it is commonly reported as the kernel with the best performance [71]), with a set of different hyperparameters, $C = [10, 1, 0.1]$ and $gamma = [0.1, 0.01, 0.001]$. These logarithmic grids are commonly tested within hyperparameter tuning, as they provide a relatively large range. Nevertheless, if best parameters lie on the boundaries of these ranges, they can be extended in subsequent analysis. The best set of hyperparameters will be determined by an exhaustive search process (using grid search[13]), which tries all possible hyperparameter combinations and evaluates the performance of the model (choosing the hyperparameter set that displays best accuracy overall).

Evidently, in order to test the performance of the model, an independent validation set is needed. For that, the method used in this work performs 5-fold cross-validation, splitting the train set into 5 folds and using a different one as the validation set for each iteration.

2-Fold Testing

The testing phase of the algorithm was designed to simulate an online classification procedure, i.e. for each sample received by the classifier, one label is assigned. This replicates the behavior a real-time classifier would need to have, in order to convey a notification every time a seizure is detected.

An additional step was added to the classification pipeline, immediately after each sample is classified by the SVM classifier, with the purpose of decreasing the algorithm's sensitivity to false positives and false negatives. It acts as a correction mechanism, based on the following constraints:

[11] Scikit-Learn. "Support Vector Machines". [Online].
[12] Scikit-Learn. "RBF SVM Parameters". [Online].
[13] With *Scikit-learn* `GridSearchCV()` function

- If a single sample is classified as background within two seizure samples, it is re-classified as seizure.

- Seizures are only classified as such, if the output classification is consistent for a period of, at least, 6 seconds[14].

The implementation of this algorithm is illustrated in Figure 5.7, in which we can perceive that it adds a small delay to the output of the classification, corresponding to the temporal constraint of 6 seconds. Furthermore, an offline version (without any temporal constraints) was also designed for testing, with the purpose of determining the added value of this approach.

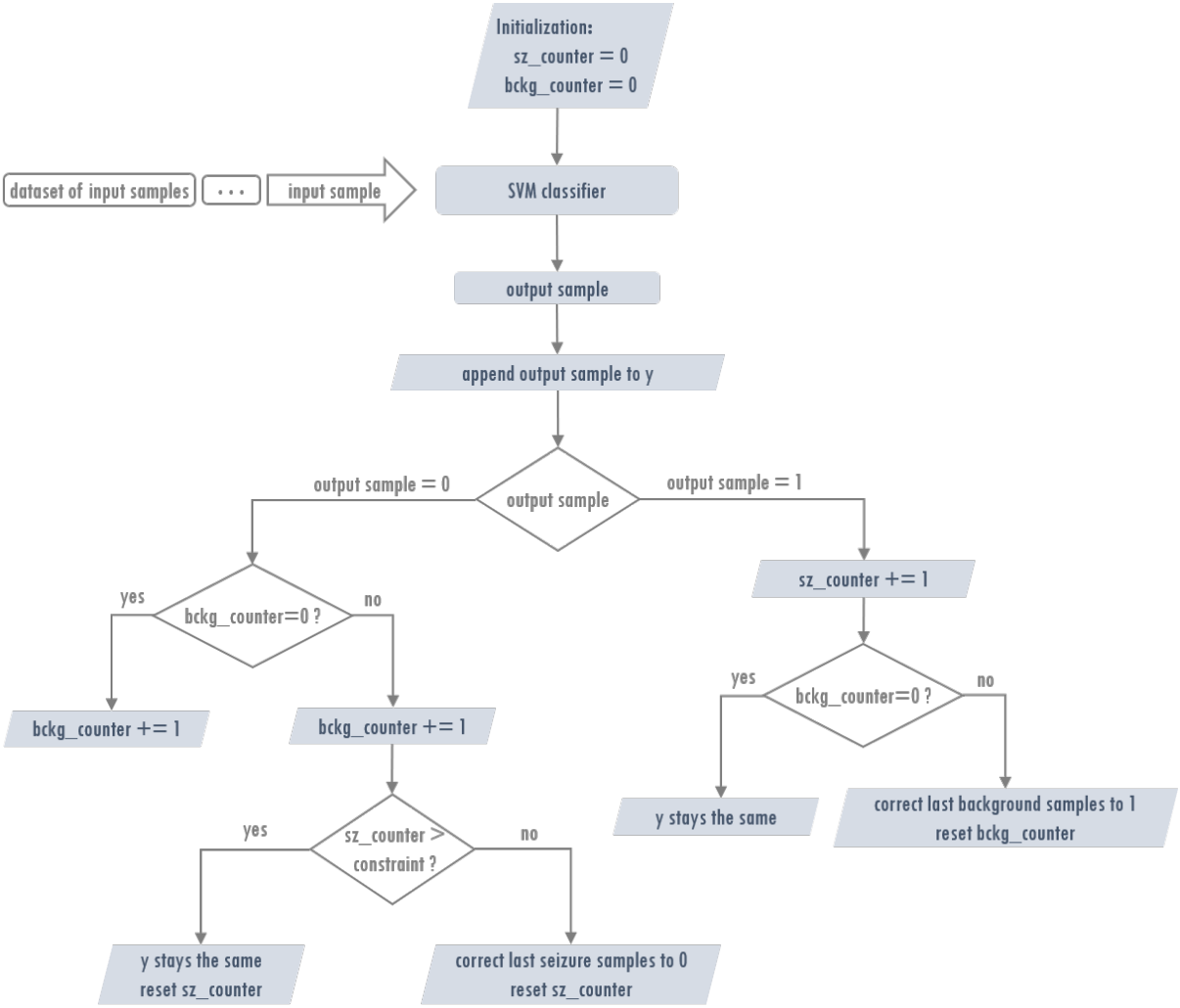


Figure 5.7: Implementation of the correction algorithm, applied within the classification pipeline, after each sample is classified by the SVM classifier. *bckg_counter*: counter that keeps track of the consecutive samples classified as 0 (background), *sz_counter*: counter that keeps track of the consecutive samples classified as 1 (seizure), *constraint*: temporal constraint corresponding to the predefined minimum duration of a seizure event.

[14]Literature reports that EEG abnormalities are required to persist for a minimum of 6-10 seconds in order to be considered a seizure. [52]

5.2 Analysis of Experimental Results

5.2.1 Best Hyperparameters

The best hyperparameters were chosen as the ones that obtained the best overall score across the 5 cross validations, performed during training. The corresponding hyperparameters are displayed in Table 5.8.

Interestingly, there does not seem be be a particular predominance of values for either parameter, as they vary greatly across montages and seizure types. However, several of theses parameters correspond to the boundaries of the logarithmic grids used in the grid search process, which might indicate the need to extend them in future analysis.

Table 5.8: Results from the training process of each model. The C and $gamma$ correspond to the parameters that obtained the best overall score over the 5 cross-validations.

	0103		**02**	
	C	$gamma$	C	$gamma$
FNSZ	10	0.1	0.1	0.1
GNSZ	0.1	0.01	1	0.1
SPSZ	1	0.001	–	–
CPSZ	1	0.01	10	0.1
ABSZ	–	–	1	0.01
TNSZ	10	0.1	–	–
TCSZ	1	0.01	1	0.1
MYSZ	–	–	10	0.01

5.2.2 Test Results

The results from the framework described above allow for three different frames of analysis: the feasibility of the proposed electrode configuration in detecting each type of seizure, independently; the added value of the correction implementation using temporal constraints; and the impact of the reference point. For simplicity and ease of visualization, only one performance metric will be presented within this discussion, nevertheless Appendix C.2 holds a more comprehensive set of metrics that can be consulted if desired. Table 5.9 shows the classification results for the test subsets, using the best hyperparameters. The values correspond to the average of F1-scores across seizure and background classes[15].

We can start by acknowledging that the online format of classification, which includes the correction algorithm, is consistently superior to its offline counterpart (except for GNSZ, SPSZ and TNSZ of montage *0103*). To validate this inference, a Wilcoxon signed-rank test for pairwise testing was adopted, under the null hypothesis that the median of the differences in performances (between offline and online) was positive against the alternative that it was negative. This statistic test does not assume normality of the data, which is crucial, considering that this would not be possible to assure with such a small sample size. The test showed that the correction algorithm did not elicit a statistically significant improvement in the performance of the classification using the subsets from montage *0103* ($w = 11.0$, $p = 0.58$), how-

[15]Recall that F1-score is a performance metric that establishes a good trade-off between sensitivity and precision.

Table 5.9: Classification results for the test subsets, using the best hyperparameters, defined above. The values correspond to the average of F1-scores, averaged over seizure and background classes (in %).

	0103		**02**	
	offline	**online**	**offline**	**online**
FNSZ	79.9	83.5	70.1	72.5
GNSZ	53.3	51.0	60.6	61.6
SPSZ	47.1	40.8	–	–
CPSZ	41.3	42.4	62.2	63.3
ABSZ	–	–	74.6	82.2
TNSZ	30.7	29.3	–	–
TCSZ	73.9	78.9	99.2	99.5
MYSZ	–	–	93.4	99.0

ever, it did demonstrate a statistically significant improvement in the subsets from montage *02* ($w = 0.0$, $p = 0.02$).

Given this, it is also interesting to note that montage *02* has overall better results than montage *0103* (except for the detection of FNSZ), which again reflects the overlap between classes in *0103*. This contributes to the initial assumption that the reference used in the montage does have the potential to significantly impact the classification, even after re-referencing. However, contrary to the reference used in *0103* (Average Reference), the reference used in *02* (Linked-Ears Reference) is largely similar to the one in EmotAI's headband (Left Ear Reference). Therefore, the fact that this montage exhibits better results is encouraging in regards to the proposed future implementations.

It is possible to recognize that, across montages and testing formats, there is a degree of consistency in the performance of each seizure type, i.e. algorithms that perform well in one of the configurations, also perform well in the others, and vice-versa. However, considering the evidence of the effect of the montage, the analysis regarding the individual seizure types will be focused on montage *02* alone. Table 5.10 shows a more comprehensive group of performance measures for montage *02* with online testing.

Table 5.10: Classification results for montage *02*, with the online testing format (including the correction algorithm).

	Accuracy %	**Sensitivity %**	**Precision %**
FNSZ	77.8	99.4	73.6
GNSZ	64.0	41.5	69.3
CPSZ	63.5	57.9	63.2
ABSZ	90.5	98.6	54.4
TCSZ	99.5	98.9	100
MYSZ	99.0	98.2	99.8

Although FNSZ was not the type of seizure with best performance, the reported results are satisfactory, with an almost perfect sensitivity to seizure events, albeit with a significant presence of false positives. The fact that this type encompasses a large category of seizures with different focalities, might induce greater variability within the seizure class, with some overlapping with the background class, and thus causing the false positives.

Similarly, ABSZ shows a large sensitivity, comparable to the one reported in [43], for the detection of absence seizures, using the same electrode configuration (sensitivity: 93.7%). However, the test results show a significantly lower precision than that of [43] (precision: 86.7%). Even though this study uses self-recorded EEG and, therefore, the results can not be straightforwardly compared, this difference in performance might suggest that this model still has potential for improvement.

Considering that GNSZ and CPSZ display considerably lower overall performances than that of the remaining types of seizures, can lead us to infer that the chosen configuration of electrodes cancels out the relevant electric patterns that would be crucial to distinguish between a seizure event and background. Although montage *02* did not have any SPSZ or TNSZ events, the results obtained with *0103* indicate the performances would likely be similar to GNSZ and CPSZ. Nevertheless, it is relevant to notice that CPSZ and SPSZ, similarly to FNSZ are very broad categories of seizures, in the sense that they can have several different focalities, which can result in contrasting detection performances for different seizure onset localizations and, hence, overall lower performance.

As anticipated during the analysis of the overlap coefficients, the best overall performance is found for TCSZ, closely followed by MYSZ, with performances comparable (and in most cases, superior) to the state of the art reported for limited channel configurations, as evidenced in Table 5.11. These results are extremely encouraging, as they provide significant evidence of the possibility of application of such a limited-channel configuration, that can be easily transposed to real-life environments, which becomes even more relevant in a type of seizure as prevalent as tonic-clonic.

Table 5.11: Table of comparison between the results of this work and previous homologous works. (*): patient-specific approach, N.S.: Not Specified. Acc: Accuracy, Sn: Sensitivity, P: Precision, Sp: Specificity.

	Seizures	**Channels**	**Acc (%)**	**Sn (%)**	**P (%)**	**Sp (%)**
This work	TCSZ MYSZ	Fp1-Fp2	99.5 99.0	98.9 98.2	100 99.8	99.2 98.3
[18] (*)	N.S.	3 channels	–	80.9	40.5	–
[42] (*)	N.S.	1 channel 2 channels	93.5 95.2	–	–	–
[43]	ABSZ	F7-FP1 F7-F3 Fp1-Fp2	–	99.1 99.1 93.7	94.8 90.2 86.7	–
[44]	GNSZ	Fp2-F8	92.7	–	–	–
[45]	N.S.	F8-T4, F7-T3	–	93.8	–	93.4
[46]	N.S.	P4-O2, C3-Cz	–	31.2	–	40.8

Chapter 6

Conclusions

6.1 Achievements

With this work we showed that it is possible to automatically detect Tonic-Clonic Seizures (TCSZ) and Myoclonic Seizures (MYSZ) using the limited channel configuration of *Fp1-Fp2*. The algorithms designed for these two seizure types achieved sensitivities of 98.9% and 98.2%, as well as precisions of 100% and 99.8%, respectively, in the task of distinguishing between non-ictal and ictal samples. These results are incredibly satisfactory, representing performances comparable (and in most cases superior) to the state-of-the-art of similar works.

Furthermore, it also suggested the possibility of extending this detection to Focal Non-Specific Seizures (FNSZ) and Absence Seizures (ABSZ), with further analysis regarding these two seizure types, as they achieved moderately adequate performances. They reported very high sensitivities of 99.4% and 98.6%, but disappointingly low precisions of 73.6% and 54.4%. These results clearly indicate the need to improve the representation of the background class, as it was typically mistaken by seizure samples. Additionally, the obtained results suggest that this particular channel configuration might not be adequate to detect Generalized Seizures (GNSZ), Complex Partial Seizures (CPSZ), Simple Partial Seizures (SPSZ) and Tonic Seizures (TNSZ), considering their overall lower performance results in all testing configurations.

By performing separate analysis for each montage, we were able to verify the impact of the reference point in the overall classification task, showing that the montage with Linked-Ears Reference (LE) has consistently better performances than the one with Average Reference (AR). Nevertheless, these results are promising, since the proposed future implementations of the detection algorithm are intended for EmotAI's headband: a limited-channel EEG acquisition wearable, with a reference point very similar to LE (i.e. Left-Ear Reference).

Regarding the autonomous system for data acquisition, *EpiBOX* operates as a standalone unit that acquires, displays and stores up to 12 different biosignal channels, simultaneously. It operates within a private wireless network, in which an MQTT broker was successfully established, guaranteeing the communication between the User Interface (UI) and the autonomous recording unit (Raspberry Pi).

A preliminary technical characterization of *EpiBOX* was performed, indicating satisfactory properties of the system, including a large efficiency in the acquisition (with only 0.3% of non-covered acquisition time), moderate memory usage and energy consumption, adequate WiFi speed to provide near-real time transfer between the acquisition device and the UI, estimated as 0.165 ms, as well as a suitable Bluetooth range to operate within an inpatient monitoring visit or home environment.

Furthermore, *EpiBOX* addresses all the issues identified in the currently available biosignal acquisition tool, as it guarantees the same acquisition flexibility (imposing no restrictions in terms of which signals to acquire, nor relating to acquisition configurations); while, at the same time, providing a simple interface, which enables the acquisition by non-technical personnel, contained within a $(82 \times 108)mm$ case.

Despite some efforts that still remain to achieve the ultimate goal of this work, this dissertation supports the potential applicability and encourages further research in the use of EmotAI's headband in the automatic detection of TCSZ, MYSZ (and potentially FNSZ and ABSZ). Furthermore, it simultaneously provides all the necessary ground work to implement the prospected autonomous system for continuous data acquisition, integrated with the obtained automatic seizure detection algorithms.

6.2 Practical Guidelines for Future Work

This dissertation belongs to a larger collaborative work, within Instituto de Telecomunicações, which intends to research and develop technological solutions to mitigate several drawbacks in the field of epilepsy. This work in particular acts as the necessary stepping stone to achieve one of the goals of providing a fully automated seizure detection system with continuous monitoring and biosignal acquisition.

With that in mind, future work should start with three angles of action: 1) carry out the usability testing phase, which should enable the detection of limitations and potential improvements of the proposed acquisition system, with the purpose of optimizing the user experience during handling of *EpiBOX*; 2) build upon the presented mobile app to create a patient version, integrated with a seizure diary, thus providing environmental context for the recorded seizures; and 3) initiate the assembly of an EEG dataset with the use of EmotAI's headband. This third task envisions the study of the applicability of the automatic detection algorithms that were developed within this dissertation to the data acquired with the wearable device. Although this work reports encouraging results using the channel *Fp1-Fp2* for the detection of TCSZ, MYSZ (and potentially FNSZ and ABSZ), performing a similar study with data acquired with the actual device is crucial to take accurate conclusions on the subject.

Furthermore, as was mentioned within this dissertation, an interesting analysis (that was unfortunately not possible to perform in this work) is the study of the impact of epoch length in the automatic detection. The author suggests the use of the present research as ground work to select a smaller set of features, in order to decrease the time necessary for feature extraction and consequently enabling this more comprehensive analysis. Other pertinent considerations include the investigation of alternative limited-channel configurations to surpass the limitations identified in this work, particularly considering the detection of GNSZ, CPSZ, SPSZ and TNSZ.

Additionally, the advantages of using multimodal approaches in seizure detection and prediction have been acknowledged in the state-of-the-art. Therefore, the author encourages the use of *EpiBOX* for simultaneous acquisition of EEG and other modalities, to further investigate the effect of such approaches in the scope of epileptic seizures.

Finally, it is impossible to not recognize the invaluable advantages of seizure prediction, i.e. anticipating when a seizure will occur, based on the dynamics of pre-ictal periods. It is to no surprise the large number of research efforts that have already been done towards the automatization of this task [6, 22, 74], even towards wearable versions of it (as extensively described in [5]). Hence, it is undeniable that this work will eventually culminate in a fully automated seizure prediction system, also with several efforts within the research team of IT being directed towards this reality. Nevertheless, all the work done in this dissertation serves either as a complementary tool to this objective (namely, *EpiBOX*) or as a background research to build useful knowledge upon.

Bibliography

[1] E. Bromfield, J. Cavazos, and J. Sirven, “Basic Mechanisms Underlying Seizures and Epilepsy,” in *An Introduction to Epilepsy*, Jan Fagerberg and David C. Mowery and Richard R. Nelson, Ed. West Hartford (CT): American Epilepsy Society, 2006, ch. 1, p. 17.

[2] Westat, *The Community Health Worker’s Training Curriculum for Epilepsy Self-Management*, Westat, 10 2017.

[3] A. H. Shoeb, “Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment,” Ph.D. dissertation, Massachusetts Institute of Technology, sep 2009.

[4] V. Shah, E. von Weltin, S. Lopez, J. McHugh, L. Veloso, M. Golmohammadi, I. Obeid, and J. Picone, “The Temple University Hospital Seizure Detection Corpus,” *Frontiers in Neuroinformatics*, vol. 12, p. 83, 01 2018.

[5] M. Abreu, A. Fred, H. Plácido da Silva, and C. Wang, “From Seizure Detection to Prediction: A Review of Wearables and Related Devices Applicable to Epilepsy via Peripheral Measurements,” Institute of Telecommunications, Tech. Rep. IT-FMCI-20190905, 03 2020.

[6] H. Agboola, C. Solebo, D. Aribike, A. Lesi, and A. Susu, “Seizure Prediction with Adaptive Feature Representation Learning,” *Journal of Neurology and Neuroscience*, vol. 10, p. 294, jan 2019.

[7] C. Guerrero-Mosquera, A. Malanda, and A. Navia-Vazquez, “EEG Signal Processing for Epilepsy,” in *Epilepsy - Histological, Electroencephalographic and Psychological Aspects*, feb 2012, vol. 3, pp. 50–58.

[8] J. Malmivuo and R. Plonsey, *Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press, New York, 10 1995.

[9] H. E. Scharfman, “The Neurobiology of Epilepsy,” *Current Neurology and Neuroscience Reports*, vol. 7, no. 4, pp. 348–354, jul 2007.

[10] J. S. Meyer and L. F. Quenzer, “Structure and Function of the Nervous System,” in *Psychopharmacology: Drugs, the Brain, and Behavior*, 2nd ed., H. de Wit, V. H. Curran, A. L. Morrow, K. Hashimoto, O. Howes, S. B. Floresco, and D. D’Souza, Eds. Sinauer Associates, 2018, ch. 2, pp. 52–56.

[11] F. E. Dudek and K. J. Staley, "The Time Course and Circuit Mechanisms of Acquired Epileptogenesis," in *Jasper's Basic Mechanisms of the Epilepsies*, 4th ed., J. L. Noebels, M. Avoli, M. A. Rogawski, R. W. Olsen, and A. V. Delgado-Escueta, Eds. Bethesda (MD): National Center for Biotechnology Information (US), 2012, pp. 595–611.

[12] D. Ochal, S. Rahman, S. Ferrell, T. Elseify, I. Obeid, and J. Picone, *The Temple University Hospital EEG Corpus: Annotation Guidelines*, The Neural Engineering Data Consortium, April 2020.

[13] E. K. St Louis, "Minimizing AED Adverse Effects: Improving Quality of Life in the Interictal State in Epilepsy Care," *Current Neuropharmacology*, vol. 7, no. 2, pp. 106–114, jun 2009.

[14] M. Tripathi and M. M. Mehendiratta, *Epilepsy Topics*. IntechOpen, 2014, ch. Role of EEG in Epilepsy, p. 15.

[15] C. Petitmengin, M. Baulac, and V. Navarro, "Seizure Anticipation: Are Neurophenomenological Approaches Able to Detect Preictal Symptoms?" *Epilepsy & Behavior : E&B*, vol. 9, no. 2, pp. 298–306, sep 2006.

[16] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG Analysis of Epilepsy: A Review," *Knowledge-Based Systems*, vol. 45, pp. 147–165, June 2013.

[17] R. M. Rangayyan, *Biomedical Signal Analysis*, 2nd ed., T. Samad, Ed. John Wiley and Sons, 2015.

[18] J. Birjandtalab, M. Baran Pouyan, D. Cogan, M. Nourani, and J. Harvey, "Automated Seizure Detection Using Limited-Channel EEG and Non-Linear Dimension Reduction," *Computers in Biology and Medicine*, vol. 82, pp. 49–58, 2017.

[19] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*. MIT Press, 2012.

[20] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[21] J. S. Marques, "Machine Learning: Evaluation and Generalization," Instituto Superior Técnico, Tech. Rep., 2017.

[22] S. M. Usman, M. Usman, and S. Fong, "Epileptic Seizures Prediction Using Machine Learning Methods," *Computational and Mathematical Methods in Medicine*, vol. 2017, p. 9074759, 2017.

[23] S. J. Russel and P. Norvig, *Artificial Intelligence: A Modern Approach*. Pearson, 2015.

[24] M. Hossin, "A Review on Evaluation Metrics for Data Classification Evaluations," *International Journal of Data Mining & Knowledge Management Process*, vol. 5, pp. 1–11, 2015.

[25] L. Guo, D. Rivero, J. A. Seoane, and A. Pazos, "Classification of EEG Signals Using Relative Wavelet Energy and Artificial Neural Networks," in *Proc. of the 1st ACM/SIGEVO Summit on Genetic and Evolutionary Computation*, 01 2009, pp. 177–184.

[26] N. K. Al-Qazzaz, S. H. Bin Mohd Ali, S. A. Ahmad, M. S. Islam, and J. Escudero, “Selection of Mother Wavelet Functions for Multi-Channel EEG Signal Analysis during a Working Memory Task,” *Sensors (Basel, Switzerland)*, vol. 15, no. 11, pp. 29 015–29 035, nov 2015.

[27] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, “Wavelet-Based EEG Processing for Computer-Aided Seizure Detection and Epilepsy Diagnosis,” *Seizure - European Journal of Epilepsy*, vol. 26, pp. 56–64, mar 2015.

[28] Bluetooth SIG Inc. Technology: Understanding Bluetooth Range. [Online]. Available: https://www.bluetooth.com/learn-about-bluetooth/bluetooth-technology/range/

[29] MQTT.org. MQTT: The Standard for IoT Messaging. [Online]. Available: https://mqtt.org/

[30] A. Banks, E. Briggs, K. Borgendale, and R. Gupta, *MQTT Version 5.0*, OASIS, march 2019.

[31] P. S. Pandher and K. K. Bhullar, “Smartphone Applications for Seizure Management,” *Health Informatics Journal*, vol. 22, pp. 209–220, 2016.

[32] 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.

[33] P. Fergus, D. Hignett, A. Hussain, D. Al-Jumeily, and K. Abdel-Aziz, “Automatic Epileptic Seizure Detection Using Scalp EEG and Advanced Artificial Intelligence Techniques,” *BioMed Research International*, vol. 2015, p. 17, 01 2015.

[34] U. Orhan, M. Hekim, and M. Ozer, “EEG Signals Classification Using the K-Means Clustering and a Multilayer Perceptron Neural Network Model,” *Expert Systems with Applications*, vol. 38, pp. 13 475–13 481, sep 2011.

[35] A. Hamad, E. Houssein, A. E. Hassanien, and A. Fahmy, “A Hybrid EEG Signals Classification Approach Based on Grey Wolf Optimizer Enhanced SVMs for Epileptic Detection,” in *Proc. of the Int’l Conf. on Advanced Intelligent Systems and Informatics 2017*, vol. 639, jan 2018, pp. 108–117.

[36] L. Guo, D. Rivero, and A. Pazos, “Epileptic Seizure Detection Using Multiwavelet Transform Based Approximate Entropy and Artificial Neural Networks,” *Journal of Neuroscience Methods*, vol. 193, no. 1, pp. 156–163, oct 2010.

[37] P. Mirzaei, G. Azemi, N. Japaridze, and B. Boashash, “Surrogate Data Test for Nonlinearity of EEG Signals: A Newborn EEG Burst Suppression Case Study,” *Digital Signal Processing*, vol. 70, pp. 30–38, 2017.

[38] C. Pradhan, S. K. Jena, S. R. Nadar, and N. Pradhan, “Higher-Order Spectrum in Understanding Nonlinearity in EEG Rhythms,” *Computational and Mathematical Methods in Medicine*, vol. 2012, p. 206857, 2012.

[39] U. R. Acharya, S. V. Sree, A. P. C. Alvin, R. Yanti, and J. SSuri, “Application of Non-Linear and

Wavelet Based Features for the Automated Identification of Epileptic EEG Signals," *International Journal of Neural Systems*, vol. 22, p. 1250002, 2012.

[40] L. Logesparan, A. J. Casson, and E. Rodriguez-Villegas, "Optimal Features for Online Seizure Detection," *Medical & Biological Engineering & Computing*, vol. 50, no. 7, pp. 659–669, 2012.

[41] A. Schulze-Bonhage, F. Sales, K. Wagner, R. Teotonio, A. Carius, A. Schelle, and M. Ihle, "Views of Patients with Epilepsy on Seizure Prediction Devices." *Epilepsy & Behavior : E&B*, vol. 18, no. 4, pp. 388–396, aug 2010.

[42] L. A. Moctezuma and M. Molinas, "EEG Channel-Selection Method for Epileptic-Seizure Classification Based on Multi-Objective Optimization," *Frontiers in Neuroscience*, vol. 14, p. 593, 2020.

[43] E. B. Petersen, J. Duun-Henriksen, A. Mazzaretto, T. W. Kjær, C. E. Thomsen, and H. B. D. Sorensen, "Generic Single-Channel Detection of Absence Seizures," in *Proc. of the 2011 Annual Int'l Conf. of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 4820–4823.

[44] S.-K. Lin, Istiqomah, L.-C. Wang, C.-Y. Lin, and H. Chiueh, "An Ultra-Low Power Smart Headband for Real-Time Epileptic Seizure Detection," *IEEE journal of Translational Engineering in Health and Medicine*, vol. 6, p. 2700410, aug 2018.

[45] D. Sopic, A. Aminifar, and D. Atienza, "e-Glass: A Wearable System for Real-Time Detection of Epileptic Seizures," in *Proc. of the 2018 IEEE Int'l Symp. on Circuits and Systems (ISCAS)*, May 2018, pp. 1–5.

[46] V. Shah, M. Golmohammadi, S. Ziyabari, E. Von Weltin, I. Obeid, and J. Picone, "Optimizing Channel Selection for Seizure Detection," in *Proc. of the 2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 2017, pp. 1–5.

[47] C. Baumgartner, J. P. Koren, and M. Rothmayer, "Automatic Computer-Based Detection of Epileptic Seizures," *Frontiers in Neurology*, vol. 9, p. 639, aug 2018.

[48] G. Regalia, F. Onorati, M. Lai, C. Caborni, and R. W. Picard, "Multimodal Wrist-Worn Devices for Seizure Detection and Advancing Research: Focus on the Empatica Wristbands," *Epilepsy Research*, vol. 153, pp. 79–82, jul 2019.

[49] T. Rukasha, S. I. Woolley, T. Kyriacou, and T. Collins, "Evaluation of Wearable Electronics for Epilepsy: A Systematic Review," *Electronics 2020*, vol. 9, no. 968, 2020.

[50] J. Pavei, R. G. Heinzen, B. Novakova, R. Walz, A. J. Serra, M. Reuber, A. Ponnusamy, and J. L. B. Marques, "Early Seizure Detection Based on Cardiac Autonomic Regulation Dynamics," *Frontiers in Physiology*, vol. 8, p. 765, oct 2017.

[51] K. Vandecasteele, T. De Cooman, Y. Gu, E. Cleeren, K. Claes, W. V. Paesschen, S. V. Huffel, and B. Hunyadi, "Automated Epileptic Seizure Detection Based on Wearable ECG and PPG in a Hospital Environment," *Sensors (Basel, Switzerland)*, vol. 17, no. 10, p. 2338, oct 2017.

[52] A. Shoeb and J. Guttag, “Application of Machine Learning To Epileptic Seizure Detection,” in *Proc. of the 27th Int’l Conf. on Machine Learning 2010*, aug 2010, pp. 975–982.

[53] P. Sirpal, A. Kassab, P. Pouliot, D. K. Nguyen, and F. Lesage, “fNIRS Improves Seizure Detection in Multimodal EEG-fNIRS Recordings,” *Journal of Biomedical Optics*, vol. 24, no. 5, pp. 1 – 9, 2019.

[54] A. J. Casson, “Wearable EEG and Beyond,” *Biomedical Engineering Letters*, vol. 9, no. 1, pp. 53–71, feb 2019.

[55] S. Debener, R. Emkes, M. De Vos, and M. Bleichner, “Unobtrusive Ambulatory EEG Using a Smartphone and Flexible Printed Electrodes Around the Ear,” *Scientific Reports*, vol. 5, no. 1, p. 16743, 2015.

[56] Y. Gu, E. Cleeren, J. Dan, K. Claes, W. Van Paesschen, S. Van Huffel, and B. Hunyadi, “Comparison Between Scalp EEG and Behind-the-Ear EEG for Development of a Wearable Seizure Detection System for Patients with Focal Epilepsy,” *Sensors (Basel, Switzerland)*, vol. 18, no. 1, p. 29, dec 2017.

[57] H. P. da Silva, A. Fred, and R. Martins, “Biosignals for Everyone,” *IEEE Pervasive Computing*, vol. 13, no. 4, pp. 64–71, 2014.

[58] D. Batista, H. P. Silva, A. Fred, C. Moreira, M. Reis, and H. A. Ferreira, “Benchmarking of the BITalino Biomedical Toolkit Against an Established Gold Standard,” *Healthcare Technology Letters*, vol. 6, pp. 32–36, 2019.

[59] K. Kutt, W. Binek, P. Misiak, G. Nalepa, and S. Bobek, “Towards the Development of Sensor Platform for Processing Physiological Data from Wearable Sensors,” in *Proc. of the 17th Int’l Conf. on Artificial Intelligence and Soft Computing 2018*, 06 2018, pp. 168–178.

[60] MathWorks. Downsampling — Aliasing. [Online]. Available: https://www.mathworks.com/help/signal/ug/downsampling-aliasing.html

[61] I. Obeid and J. Picone, “The Temple University Hospital EEG Data Corpus,” *Frontiers in Neuroscience*, vol. 10, p. 196, 2016.

[62] S. Ferrell, , V. Mathew, M. Refford, V. Tchiong, T. Ahsan, I. Obeid, and J. Picone, *The Temple University Hospital EEG Corpus: Electrode Location and Channel Labels*, The Neural Engineering Data Consortium, July 2019.

[63] S. López, A. Gross, S. Yang, M. Golmohammadi, I. Obeid, and J. Picone, “An Analysis of Two Common Reference Points for EEGs,” *IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, vol. 2016, p. 10.1109/SPMB.2016.7846854, dec 2016.

[64] A. Widmann, E. Schröger, and B. Maess, “Digital Filter Design for Electrophysiological Data – A Practical Approach,” *Journal of Neuroscience Methods*, vol. 250, pp. 34 – 46, 2015.

[65] P. Lio and Y. Song, “A New Approach for Epileptic Seizure Detection: Sample Entropy Based Feature Extraction and Extreme Learning Machine,” *Journal of Biomedical Science and Engineering*, vol. 3, pp. 556–567, 01 2010.

[66] 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, pp. H2039–49, jun 2000.

[67] A. Delgado-Bonal and A. Marshak, “Approximate Entropy and Sample Entropy: A Comprehensive Tutorial,” *Entropy*, vol. 21, no. 6, 2019.

[68] T. N. J. Kawe, S. M. Shadli, and N. McNaughton, “Higuchi’s Fractal Dimension, but not Frontal or Posterior Alpha Asymmetry, Predicts PID-5 Anxiousness More Than Depressivity,” *Scientific Reports*, vol. 9, no. 1, p. 19666, 2019.

[69] S. Kesić and S. Spasic, “Application of Higuchi’s Fractal Dimension from Basic to Clinical Neurophysiology: A Review,” *Computer Methods and Programs in Biomedicine*, vol. 133, pp. 55–70, 05 2016.

[70] S. Chandrasekaran, S. Poomalai, B. Saminathan, S. Suthanthiravel, K. Sundaram, and F. F. Abdul Hakkim, “An Investigation on the Relationship Between the Hurst Exponent and the Predictability of a Rainfall Time Series,” *Meteorological Applications*, vol. 26, no. 3, pp. 511–519, 2019.

[71] U. R. Acharya, R. Yanti, G. Swapna, V. S. Sree, R. J. Martis, and J. S. Suri, “Automated Diagnosis of Epileptic Electroencephalogram Using Independent Component Analysis and Discrete Wavelet Transform for Different Electroencephalogram Durations,” *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of Engineering in Medicine*, vol. 227, no. 3, pp. 234–244, mar 2013.

[72] B. Butcher and B. J. Smith, “Feature Engineering and Selection: A Practical Approach for Predictive Models,” *The American Statistician*, vol. 74, no. 3, pp. 308–309, 2020.

[73] A. Subasi and M. Ismail Gursoy, “EEG Signal Classification Using PCA, ICA, LDA and Support Vector Machines,” *Expert Systems with Applications*, vol. 37, no. 12, pp. 8659 – 8666, 2010.

[74] 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, jul 2018.

Appendix A

EpiBOX

A.1 Setup

One of the most relevant features of *EpiBOX* is its portability and practicality to implement both in a hospital or home environment. In fact, this was one of the issues found in the previous acquisition setup, which was crucial to address. No image was available for the previous setup in the hospital environment, however one can easily imagine the inconvenience of having a computer on top of a stand, in the middle of a corridor of the inpatient ward. Conversely, Figure A.1 simulates the flexible setup for *EpiBOX*, which can be discretely placed even within the patient's room.

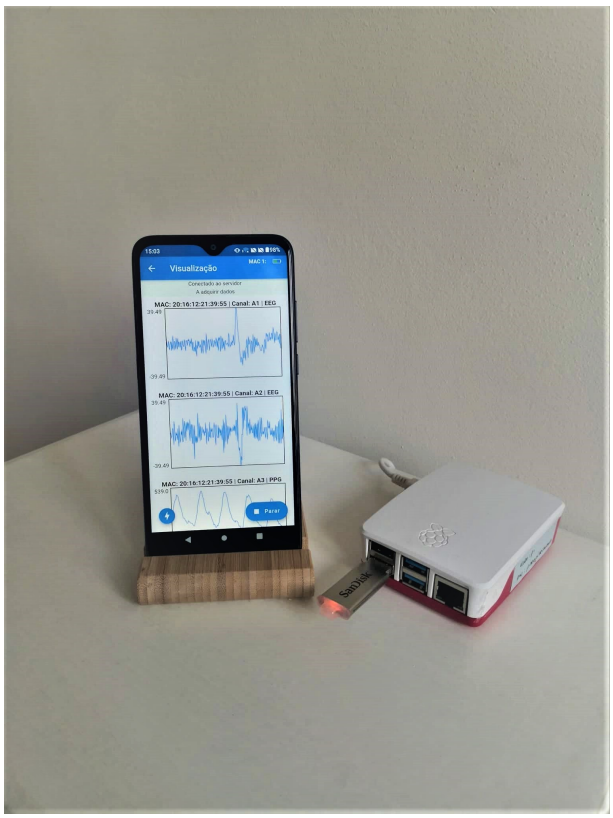


Figure A.1: Acquisition setup with *EpiBOX*.

Considering the proposed future implementations of *EpiBOX* with EmotAI's headband, Figures A.2 (a) and (b) illustrate the use of the fit of the wearable device during acquisition. Figure A.2 (c) allows to see the interior of the headband, including one of the pairs of dry electrodes positioned to acquire the channels corresponding to *Fp1* and *Fp2*, as well as the battery and BITalino board.

(a) (b)

(c)

Figure A.2: EmotAI's EEG headband. (a) and (b) highlight of the positioning of the both pairs of electrodes and reference point; (c) interior of the device.

A.2 Mobile App

The mobile app that acts as the User Interface with *EpiBOX* was designed to provide maximum control over the acquisition process, while ensuring a simple and easy to follow procedure. As such, it presents the whole process as set of one-time tasks that must be performed consecutively in an exclusive order, which is established by the step sequence in the Homepage (Figure A.3(a)). The connection to Raspberry Pi (RPi) is initiated in the Connectivity page (Figure A.3(b)). The following step implies the selection of the MAC addresses of the BITalino device(s), in the Device selection page (Figure A.3(c)), with a minimum of 1 and a maximum of 2 devices. For practical purposes, the text fields are initially set to the default MAC addresses, which can be altered either by manual typing or by QR code scanning. The final step before initiating acquisition is the selection of acquisition parameters in the Configuration page (Figure A.3(d)). This allows the flexibility of configuring the most relevant set of parameters, namely the storing folder (which can be in a hard-drive or internal storage of the RPi), sampling frequency, which channels to acquire from both BITalino devices and the corresponding sensors, if desired. Just like in the previous set, all configurations have default values, to enable a faster procedure when possible.

Finally, the acquisition can be initiated and the Visualization page is enabled (Figure A.4(a)), where it is possible to visualize the signals being acquired in near-real time. This is a scrollable page, which displays up to 12 signals, depending on the number of channels being acquired, as well as the battery of the devices in acquisition (this feature is complemented by a push notification whenever one of the batteries is at 10% or below). Moreover, during the recording process, the app allows for manual annotations in order to facilitate post-processing of the data. This is performed in the Annotation page (Figure A.4(b)), which has a list of pre-defined events (e.g. change of batteries or equipment misplacement) and a text field to add a new one, as well as the possibility of adjusting the timestamp of the event. All the annotation are stored in the RPi, along with the recording file, in a standardized, computationally-readable format.

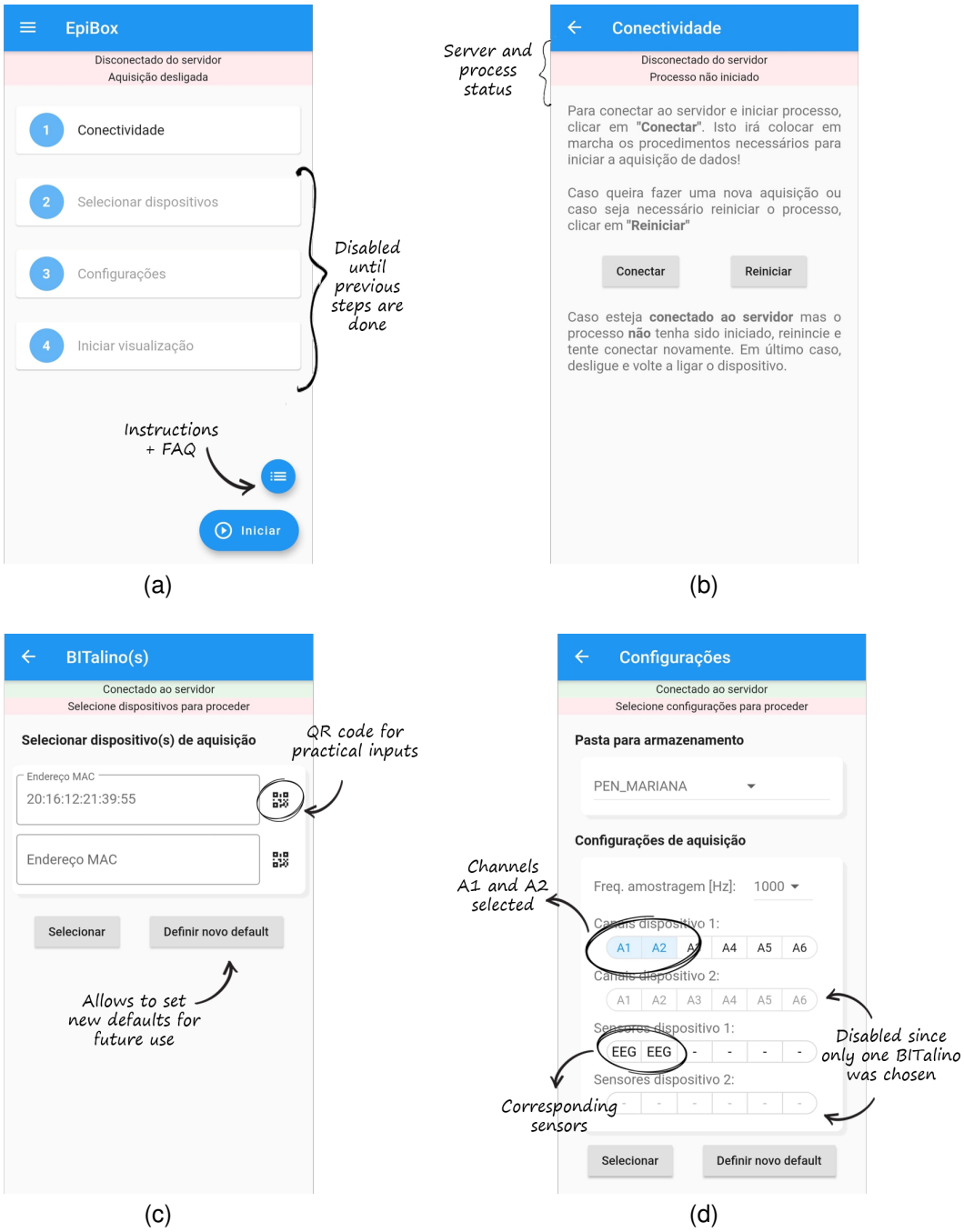


Figure A.3: Print screens from the mobile app, namely of the (a) Homepage, (b) Connectivity page, (c) Device selection page and (d) Configuration page. Some annotations are also displayed for better visualization.

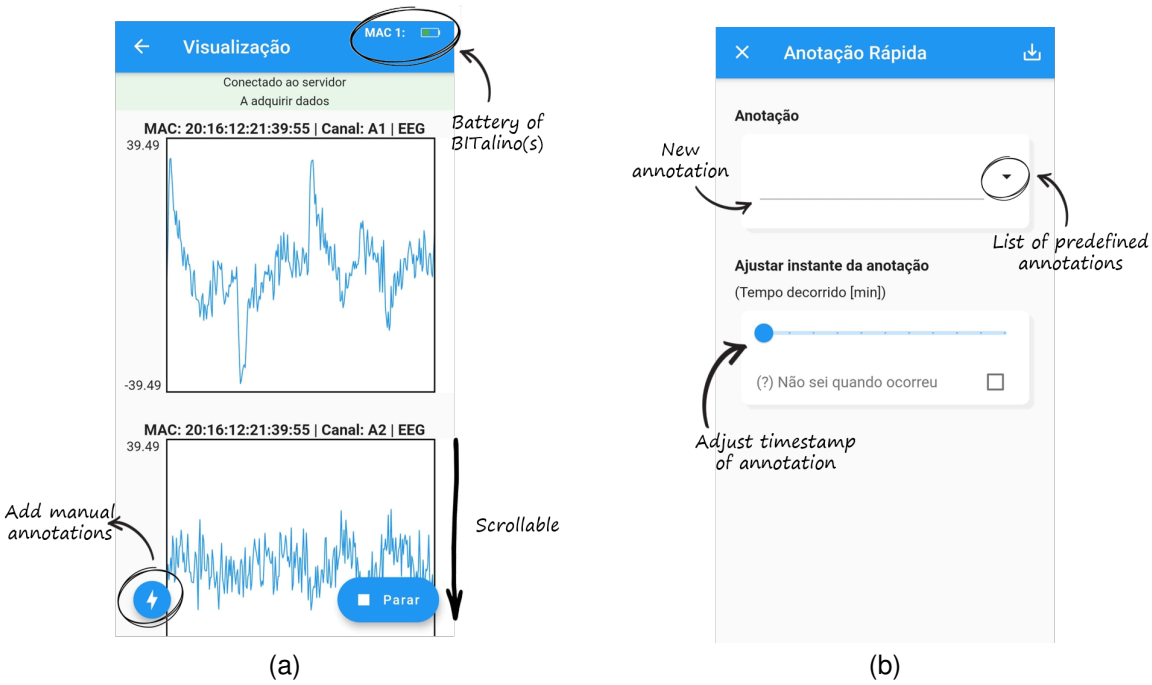


Figure A.4: Print screens from the mobile app, namely of the (a) Visualization page, (b) Annotation page. Some annotations are also displayed for better illustration of the operation of the mobile app.

A.3 File Formats

The data acquired with *EpiBOX* is saved in a *.txt* file named with the date and time of the beginning of the recording process (*yy-MM-dd HH-mm-ss*), in the corresponding folder of the patient. This file includes a header which contains all the information needed to interpret the data, as described in Table A.1, in the format of a dictionary which elicits all these properties for each of the acquired BITalino devices.

Table A.1: Properties present in the header, for each of the acquired BITalino devices.

Property	Type	Description
date	`String`	Date of the acquisition (*yy-MM-dd*).
time	`String`	Time of the acquisition (*HH:mm:ss.s*).
device connection	`String`	MAC address of BITalino.
column	`List<String>`	Column labels. Always starts with *"nSeq"* and is followed by the analog channels that BITalino acquired.
channels	`List<int>`	Analog channels that BITalino acquired.
sampling rate	`float`	Frequency of sampling used in acquisition.
sensor	`List<String>`	Specifies if the signal is in ADC or in physical units (for each channel).
label	`List<String>`	Label of the sensor, as defined in the mobile app (for each channel).
resolution	`List<int>`	Resolution of each column.

The actual content of the file is illustrated in Table A.2 and sums up to $\#channels + 2$ columns, which correspond to the data acquired by each of the channels in acquisition, as well as a sequence number per BITalino (identified as *"nSeq"*, it is an additional element provided by BITalino with each sample, that can be used to detect issues with the acquisition, i.e. if the sequence, which overflows at 15, is not

correct, there are missing samples).

Table A.2: Illustration of the structure of the data saved in the acquisition *.txt* file. Example in which the acquisition is performed with 2 BITalino devices: one acquiring channels [A1,A2,A4] and the other [A1,A2]. Note $s_{i(Aj,BITl)}$ corresponds to sample i acquired from channel j of BITalino l.

0	$s_{1(A1,BIT1)}$	$s_{1(A2,BIT1)}$	$s_{1(A4,BIT1)}$	0	$s_{1(A1,BIT2)}$	$s_{1(A2,BIT2)}$
1	$s_{2(A1,BIT1)}$	$s_{2(A2,BIT1)}$	$s_{2(A4,BIT1)}$	1	$s_{2(A1,BIT2)}$	$s_{2(A2,BIT2)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
15	$s_{16(A1,BIT1)}$	$s_{16(A2,BIT1)}$	$s_{16(A4,BIT1)}$	15	$s_{16(A1,BIT2)}$	$s_{16(A2,BIT2)}$
0	$s_{17(A1,BIT1)}$	$s_{17(A2,BIT1)}$	$s_{17(A4,BIT1)}$	0	$s_{17(A1,BIT2)}$	$s_{17(A2,BIT2)}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

In addition to the acquisition file, another *.txt* file is saved per session: the annotation file. This file contains only two columns, one corresponding to the event label (e.g. battery change) and the other to the timestamp, defined in the mobile app (*HH:mm:ss*, if known, `null`, otherwise).

A.4 MQTT Messages

In the context of the present work, MQTT is the protocol that establishes the communication between the RPi and the mobile app. It allows the control of the RPi by the user, through a pre-defined set of messages. As was described in Section 2.3.2, an MQTT message packet is composed by a fixed header with 2 bytes; a variable header which, for publishing a message, corresponds to 2 bytes as well; and a payload of variable length, corresponding to the actual message. In *EpiBOX*, the payload always corresponds to a list with a *command* as the first element (which tells the mobile app and RPi what the message is about) and the *content* as the remaining element(s), e.g.:

['DEFAULT MAC', 'xx:xx:xx:xx:xx:xx', 'xx:xx:xx:xx:xx:xx']

Table A.3 describes the pre-defined set of messages exchanged between the mobile app and the RPi, providing the command, content type, maximum packet size (computed as described above) and content description.

Table A.3: Complete set of messages exchanged between the mobile app and the RPi. (*): some commands do not have a set maximum packet size, as they depend on the user input; in such cases, the format used to compute the size is provided as an example.

Command	Type	Max Size [bytes]	Content Description
TIME	`String`	33	Current time (*'yy-MM-dd HH:mm:ss'*) to update RPi, which has no connection to the internet.
DEFAULT MAC	`String`	59	Two strings corresponding to the default MAC addresses (*'xx:xx:xx:xx:xx:xx'* or empty String if not defined).
DRIVES*	`List<String>`	50	List of the available external drives + internal storage (e.g. PEN_MARIANA).
DEFAULT CONFIG*	`List<String, int,List>`	369	List with default drive (e.g. 'PEN_MARIANA'), default sampling frequency, and list with default channels/sensors.
NEW MAC	`dict`	71	Dictionary with keys *'MAC1'* and *'MAC2'*, whose values correspond to the new default MAC addresses.
USE MAC	`dict`	71	Format as NEW MAC, but whose values correspond to the MAC addresses to use.
ID*	`String`	26	Patient ID (e.g. 'xxxxxxxx_xxxx').
FOLDER*	`String`	28	Location to store acquired data (e.g. 'PEN_MARIANA').
FS	`int`	15	Sampling frequency.
CHANNELS	`List<List<String>>`	367	List of lists with the MAC address, channel and sensor, for each channel to acquire.
NEW CONFIG DEFAULT*	`List<String, int,List>`	398	List with new defaults (each with the formats considered above).
START	–	13	No content. Tells the RPi to start the acquisition.
RESTART	–	15	No content. Tells the RPi to restart the whole process.
INTERRUPT	–	17	No content. Tells the RPi to stop the acquisition and save files.
ANNOTATION*	`List<String>`	53	List with the event (e.g. 'Mudança de bateria') and its timestamp (*'HH:mm:ss'* or *'null'* if not defined).
BATTERY	`dict`	48	Dictionary with keys corresponding to the MAC addresses in use, whose values correspond to the battery of each BITalino (in ADC).
DATA	`List<List<float>>, List<int>, List<String>`	743	Three lists, one with the data acquired from each of the channels, the second with the acquired channels and the third with the corresponding sensors.
ACQUISITION ON	–	19	No content. Tells the mobile app the acquisition is ongoing.
RECONNECTING	–	20	No content. Tells the mobile app the acquisition was interrupted for some reason and is trying to resume.
TIMEOUT	`String`	35	Whenever the process of resuming the acquisition is taking too long, the MAC address that is (possibly) causing the issue is sent.
STOPPED	–	15	No content. Tells the mobile app the acquisition has stopped and the files were saved.

Appendix B

Usability Questionnaire

This questionnaire was designed to evaluate the whole system of *EpiBOX*, from the convenience of the mobile app to the reliability of the acquisition setup, and has two main goals:

1. Assess the usability of *EpiBOX* by medical staff, without any assumption on their proficiency in the use of modern technology; and consequently evaluate the potential of *EpiBOX* to be integrated in a normal clinical context, as a collaboration tool between researchers and clinical staff.
2. Identify potential improvements and adaptations needed for an optimal operation and usability of the system.

It is divided into two independent sections, each conceived to answer one of the goals. The first section is based on the widely-used *System Usability Scale*, which provides a reliable, industry-standard tool to evaluate the ease of use of a large variety of products and services. The second section aims at identifying potential technical problems on a real-world application environment (instead of a development environment), as well as system design flaws that provide a non-optimal experience for the user and should therefore be revised. The Portuguese version of the questionnaire can be consulted below.

NOME: _____ *INÍCIO DA SESSÃO:* ___/___/_____ ___:___
(DATA) (HORA)

QUESTIONÁRIO DE USABILIDADE - *EPIBOX*

Para cada uma das seguintes afirmações, marque a caixa que melhor descreve a sua experiência com sistema ***EpiBox***, relativamente à interação com **todo o sistema**.

Escala de Usabilidade de Sistemas (SUS)		Discordo fortemente					Concordo plenamente
1.	Acho que utilizaria o sistema *EpiBox* frequentemente.	☐	☐	☐	☐	☐	☐
2.	Achei a interação desnecessariamente complexa.	☐	☐	☐	☐	☐	☐
3.	Achei fácil interagir com o *EpiBox*.	☐	☐	☐	☐	☐	☐
4.	Penso que necessitaria de suporte técnico para conseguir utilizar o *EpiBox* (para além da sessão de esclarecimento inicial).	☐	☐	☐	☐	☐	☐
5.	Achei que as diversas funções do *EpiBox* estavam muito bem integradas.	☐	☐	☐	☐	☐	☐
6.	Achei que havia muita inconsistência no *EpiBox*.	☐	☐	☐	☐	☐	☐
7.	Penso que a maioria das pessoas conseguiriam aprender a interagir com o *EpiBox* rapidamente.	☐	☐	☐	☐	☐	☐
8.	Achei que o *EpiBox* era muito desconfortável de utilizar.	☐	☐	☐	☐	☐	☐
9.	Senti-me muito confiante a usar o *EpiBox*.	☐	☐	☐	☐	☐	☐
10.	Necessitei de aprender muitas coisas antes de conseguir utilizar o *EpiBox*.	☐	☐	☐	☐	☐	☐

Para cada uma das seguintes afirmações, marque a caixa que melhor descreve a sua experiência com sistema ***EpiBox***, relativamente à interação com a **aplicação móvel** (App).

Inicialização		NA	Discordo fortemente					Concordo plenamente
1.	A forma como a App está desenhada torna o fluxo de utilização intuitivo.	☐	☐	☐	☐	☐	☐	☐
2.	Foi fácil realizar os passos necessários para iniciar a aquisição.	☐	☐	☐	☐	☐	☐	☐
3.	Foi fácil realizar os passos necessários para adicionar uma anotação.	☐	☐	☐	☐	☐	☐	☐
4.	Foi fácil realizar os passos necessários para terminar a aquisição.	☐	☐	☐	☐	☐	☐	☐
5.	As instruções tinham todas as informações necessárias para iniciar a aquisição.	☐	☐	☐	☐	☐	☐	☐
6.	As instruções tinham todas as informações necessárias para adicionar uma anotação.	☐	☐	☐	☐	☐	☐	☐
7.	As instruções tinham todas as informações necessárias para terminar a aquisição.	☐	☐	☐	☐	☐	☐	☐

8.	A App demorou mais tempo do que o esperado para iniciar a aquisição.	☐	☐	☐	☐	☐	☐	☐
Visualização		NA	Discordo fortemente					Concordo plenamente
9.	A visualização dos sinais foi sempre fluida (sem paragens bruscas).	☐	☐	☐	☐	☐	☐	☐
Resolução de problemas		NA	Discordo fortemente					Concordo plenamente
10.	A secção *RESOLUÇÃO DE PROBLEMAS,* das instruções, resolveu todas as dúvidas que me surgiram na sessão de aquisição.	☐	☐	☐	☐	☐	☐	☐
11.	Quando me deparei com um erro/problema, foi fácil de resolver.	☐	☐	☐	☐	☐	☐	☐
12.	Independentemente do problema, ao reiniciar a conexão, deixei de ter esse problema.	☐	☐	☐	☐	☐	☐	☐
13.	Independentemente do problema, ao desligar e voltar a ligar a ***EpiBox***, deixei de ter esse problema.	☐	☐	☐	☐	☐	☐	☐

Para cada uma das seguintes afirmações, marque a caixa correspondente à sua experiência nesta sessão, ainda relativamente à interação com a **aplicação móvel** (App).

Problemas – Inicialização		NA	V	F
1.	Não consegui encontrar a rede WiFi nos primeiros 5 min.	☐	☐	☐
2.	Não foi possível encontrar a rede WiFi nunca.	☐	☐	☐
3.	Não consegui conectar ao servidor à primeira tentativa.	☐	☐	☐
4.	Não foi possível conectar ao servidor nunca.	☐	☐	☐
5.	Tive dificuldades a introduzir um ou mais dados para a aquisição (dispositivos, pasta de armazenamento, frequência de amostragem ou canais / sensores a adquirir).	☐	☐	☐
Problemas – Durante a aquisição		NA	V	F
6.	Depois de trocar a(s) bateria(s), a aquisição não retomou sozinha.	☐	☐	☐
7.	Face aos canais/sensores escolhidos, faltaram gráficos de visualização.	☐	☐	☐

Agradeço a sua participação,
continuação de um excelente trabalho!

Ana Sofia Carmo

Appendix C

Seizure Detection

C.1 Train Results

For each seizure type and for each montage (*0103* and *02*), an SVM classifier with RBF kernel was trained on the train subset. In order to tune the hyperparameters - regression parameter (C) and kernel coefficient ($gamma$) - a procedure of grid search was applied (using *sklearn* `GridSearchCV`), with 5-fold cross-validation, with $C = [10, 1, 0.1]$ and $gamma = [0.1, 0.01, 0.001]$, resulting in 9 candidates for each. Table C.1 presents the chosen hyperparameters, according to the best overall accuracy score across the 5 cross validations, as well as the respective average performance values across the 5 cross validations.

Table C.1: Results from the training process of each model. The C and $gamma$ correspond to the parameters that obtained the best overall score over the 5 cross-validations and Acc (%) corresponds to the corresponding average accuracy values across the 5 cross validations.

	0103			**02**		
	C	$gamma$	**Acc (%)**	C	$gamma$	**Acc (%)**
FNSZ	10	0.1	73.4 $\pm$ 3.6	0.1	0.1	74.0 $\pm$ 1.8
GNSZ	0.1	0.01	77.0 $\pm$ 4.6	1	0.1	87.8 $\pm$ 5.3
SPSZ	1	0.001	74.4 $\pm$ 18.6	–	–	–
CPSZ	1	0.01	73.0 $\pm$ 12.7	10	0.1	63.1 $\pm$ 6.8
ABSZ	–	–	–	1	0.01	98.2 $\pm$ 2.2
TNSZ	10	0.1	73.3 $\pm$ 5.4	–	–	–
TCSZ	1	0.01	91.8 $\pm$ 11.5	1	0.1	95.2 $\pm$ 1.5
MYSZ	–	–	–	10	0.01	96.8 $\pm$ 1.1

C.2 Test Results

After training each model on the corresponding train datasets, all 12 models (corresponding to each seizure type and montage) were tested with two different approaches: 1) An offline testing in which each sample of the test dataset was classified independently; and 2) An online testing in which there was assumed a temporal dependency between consecutive samples and to which was applied a correction algorithm. The results obtained for both testing techniques are displayed in Table C.2 for montage *0103* and in Table C.3 for montage *02*. The results are a set of three performance metrics (Accuracy, Sensitivity and Precision) that complement the averaged F1-score shared in Section 5.2.

Table C.2: Classification results for the test subsets for montage *0103*.

	0103					
	Accuracy [%]		**Sensitivity [%]**		**Precision [%]**	
	offline	**online**	**offline**	**online**	**offline**	**online**
FNSZ	82.2	84.6	95.2	96.7	78.1	70.0
GNSZ	57.9	57.7	89.4	94.9	54.8	54.4
SPSZ	62.4	65.9	12.5	1.2	36.0	60.0
CPSZ	42.5	44.5	32.6	30.0	32.8	33.3
TNSZ	27.2	24.8	87.0	82.6	22.4	21.2
TCSZ	74.5	80.2	64.6	61.2	76.2	93.0

Table C.3: Classification results for the test subsets for montage *02*.

	02					
	Accuracy [%]		**Sensitivity [%]**		**Precision [%]**	
	offline	**online**	**offline**	**online**	**offline**	**online**
FNSZ	76.1	77.8	98.8	99.4	72.4	73.6
GNSZ	62.4	64.0	43.7	41.5	64.4	69.3
CPSZ	62.2	63.5	60.6	57.9	60.8	63.2
ABSZ	84.4	90.5	98.6	98.6	41.9	54.4
TCSZ	99.2	99.5	98.7	98.9	99.4	100
MYSZ	93.4	99.0	92.7	98.2	93.8	99.8