FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

NeuroMov: Multimodal approach for epileptic seizure detection and prediction

Paulo Manuel de Carvalho Branco Maia

DISSERTAÇÃO

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Mestrado Integrado em Bioengenharia - Engenharia Biomédica

Supervisor: João Paulo Trigueiros da Silva Cunha, PhD.

July 17, 2019

NeuroMov: Multimodal approach for epileptic seizure detection and prediction

Paulo Manuel de Carvalho Branco Maia

Mestrado Integrado em Bioengenharia - Engenharia Biomédica

July 17, 2019

Resumo

A epilepsia é uma das doenças neurológicas mais comuns, afetando entre 0.5 e 1% da população mundial. É caracterizada por uma predisposição de gerar crises epiléticas, uma ocorrência transiente e temporária de sinais ou sintomas que são consequentes de uma atividade neuronal anormal no cérebro.

Pacientes com epilepsia ativa podem sofrer de consequências severas resultantes das crises, como lesões ou morte. Os sistemas para deteção objetiva de crises podem ter várias funções, tais como melhorar o seguimento de crises (que, de momento, é muito subjetivo) e alertar um cuidador. Para além da deteção de crises, a sua previsão antes do seu início permitiria que os pacientes tomassem precauções contra ferimentos, por quedas e outros acidentes, e incentivar o desenvolvimento de tratamentos que controlassem a crise futura, como neuroestimuladores.

Existem vários dispositivos para detetar alterações fisiológicas ocorrentes em crises, tais como o eletroencefalograma (EEG) e o eletrocardiograma (ECG). Os dados adquiridos através destes sensores podem ajudar a distinguir o período ictal (durante a crise), pré-ictal (antes da crise, cuja duração é específica do doente) e interictal (entre crises). Dependendo do tipo de crise, diferentes alterações são vistas em cada sinal.

Nos últimos anos, várias metodologias para deteção e previsão de crises com recurso a *machine learning* e *deep learning* têm vindo a ser desenvolvidas. No entanto, muitos destes algoritmos ainda têm taxas de falsos positivos demasiado altas. Nesta dissertação, exploramos a possibilidade da melhoria da performance dos algoritmos com com uma abordagem multimodal, que combine dois ou mais biosinais de acordo com as necessidades de cada paciente.

Dados de EEG e ECG de 8 doentes epiléticos, adquiridos numa Unidade de Monitorização de Epilepsia, com diversos tempos de gravação e tipos de crises foram usados para criar um classificador multimodal para deteção (período ictal) e previsão (período interictal) de crises, específico ao doente. Um total de cerca de 110h de dados foram analisados. Foi utilizada uma metodologia *leave one seizure out cross validation*, agrupando os dados contendo o período antes de uma crise e o período da crise. Após testar alguns classificadores, uma arquitetura de *deep learning* (baseada em LSTMs) aplicada diretamente aos dados em bruto foi selecionada como o melhor classificador.

Os resultados dos classificadores separados de EEG e ECG foram combinados usando operadores lógicos - "AND" para o período ictal e "OU" para o pre-ictal. Em média, a sensibilidade da deteção diminuiu ligeiramente com a abordagem multimodal, mas levou a menos falsos positivos comparado com a unimodal. Por outro lado, o comportamento contrário é visto para a previsão.

Os resultados mostram potencial para desenvolvimento de uma abordagem multimodal específica ao doente para uma deteção e previsão de crises clinicamente utilizável, com melhorias em comparação às abordagens unimodais.

Abstract

Epilepsy is one of the most common neurological disorders, affecting from 0.5 to 1% of the world population. It is characterized by a predisposition to generate epileptic seizures, a transient and temporary occurrence of signs/symptoms which are consequence of an abnormal neuronal activity in the brain.

Patients with active epilepsy may suffer from severe consequence from seizures, such as injury or death. Objective seizure detection systems could have several functions, such as improving seizure tracking (which is as of now very subjective) or alert a caregiver. Besides seizure detection, accurately predicting seizures before their onset would enable patients to take precautions against injury, and could contribute for more effective development of treatments for controlling the upcoming seizure, such as closed-loop neurostimulators.

Several devices exist for detecting physiological changes occurring in seizures, such as the electroencephalogram (EEG) and electrocardiogram (ECG). Data acquired from theses sensors can help distinguish the ictal period (during a seizure), pre-ictal (before a seizure, with patient-dependent duration) and the interictal period (between seizures). Depending on the seizure type, different alterations are seen in each signal and period.

In recent years, numerous methodologies for seizure detection and prediction using machine learning and deep learning have been developed. However, many algorithms still have very high false positive rates, which leads to its unusability in routine clinical and ambulatory scenarios. In this dissertation, we explore the postulation that these algorithms' performance may be improved with a multimodal approach, combining two or more biosignals according to each patient's needs.

EEG and ECG data from 8 epileptic patients, acquired in Epilepsy Monitoring Units, with diverse recording length and seizure types was used for creating a patient-specific multimodal seizure detector (ictal state) and predictor (pre-ictal state). A total of around 110h of data was analyzed. A leave one seizure out cross validation was performed, grouping data containing the period before a seizure and the seizure period. After testing a few classifiers, a deep learning architecture (based on LSTM) applied directly on raw data was selected as the best.

Outputs from separate EEG and ECG classifications were combined using logical operators - "AND" for the ictal period and "OR" for the pre-ictal period. On average, detection sensitivity decreases slightly with the multimodal approach, but with less false positives compared to unimodal approaches. On the other hand, the opposite behavior is seen for prediction.

The obtained results show potential for developing a multimodal patient-specific approach for usable seizure detection and prediction, with improvements compared to unimodal approaches.

Agradecimentos

E assim termino os cinco anos mais ricos da minha vida. Neste percurso, muita gente contribuiu para o meu desenvolvimento pessoal, às quais gostaria de começar por mostrar o meu agradecimento através desta mensagem. Apesar disto, imortalizar os vossos nomes neste pedaço de texto é insuficiente para agradecer a vossa influência na minha vida.

A toda a gente que participou no NEB, sem dúvida o maior desafio deste ano. Conseguir conciliar a tese com constantes momentos de tomada de decisão e resolução de imprevistos foi uma experiência que me ajudou a crescer bastante. Um agradecimento especial ao Abel e ao Carlos, pela confiança dada através do convite para a entrada no NEB há dois anos, transformando completamente a minha vida académica.

A todo o BRAIN, mas em especial ao Duarte e à Susana pelo apoio constante na escrita dos projetos H2020, ao Hugo, por me ter integrado no grupo e ajudado a desenvolver o "bichinho" da investigação, à Maria do Carmo e à Elodie por toda a ajuda. Agradeço ainda ao Professor João Paulo Cunha, por todas as discussões, pelo incentivo a ir mais longe, pela liberdade de escolher o meu próprio caminho e participar em projetos desafiantes, e pela orientação ao longo destes anos.

A todos os amigos que fiz na FEUP estes anos. Um agradecimento especial ao grupo que convive comigo constantemente: Rajah, Cris, Varejão, Lena, Leo, Maggie, Nina, Mariana Campos, Mariana Dias, Marina, Patrícia Ribeiro, Patrícia Rocha, Alex, Bruno Pinto, Joana Rodrigues, Nuno, Vitor Minhoto, Tiago, Gabriel, Diogo Martins, Joana Dias, Filipa, Rita Gouveia. Ao metaleiro, Luís Miguel, o "grande macho". Àqueles que vejo menos, mas que continuam sempre presentes: Eduardo, Zé. Aos MEBianos: Ju, Bruno, Tomé e Luís.

Aos meus *e-amigos*, o Nuno e o Eduardo, que me acompanham à distância há mais de 8 anos.

Aos meus conterrâneos, que ainda hoje me acompanham: Rui, Nando, Renas (o colega), Clara, Fábio, Bino, Helena João e Helena Xavier - obrigado!

À minha família, por me ter aturado em todos os momentos fáceis e difíceis ao longo destes cinco anos. Todos vocês foram um excelente modelo para mim, e lutaram para que terminasse esta etapa com sucesso.

"It's better to have it and not need it, than to need it and not have it"

Louis Michael Fields

Contents

List of Figures

List of Tables

Abreviaturas e Símbolos

AED	Anti Epileptic Drugs
AES	American Epilepsy Society
AI	Artificial Intelligence
ANN	Artificial Neural Network
BPM	Beats Per Minute
CAM	Class Activation Map
CNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform
DL	Detection Latency
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
EDA	Electrodermal Activity
EEG	Electroencephalogram
EMU	Epilepsy Monitoring Unit
GPU	Graphics Processing Unit
HMM	Hidden Markov Model
HR	Heart Rate
HRV	Heartrate variability
ILAE	International League Against Epilepsy
LSTM	Long Short-term Memory
ML	Machine Learning
MLP	Multilayer Perceptron
MVC	Maximum voluntary contraction
NINDS	National Institute of Neurological Disorders and Stroke
PCA	Principal Component Analysis
PT	Prediction Time
RNN	Recurrent Neural Network
SOP	Seizure Occurrence Period
SPH	Seizure Prediction Horizon
SSDA	Stacked Sparse Denoising Autoencoder
STFT	Short-time Fourier Transform
SVM	Support Vector Machine
TCS	Tonic-Clonic Seizure
TS	Tonic Seizure
TUH	Temple University Hospital
sEMG	Surface Electromyogram

Chapter 1

Introduction

1.1 Epilepsy

Epilepsy is one of the most common neurological disorders in the world (1). Just in Europe, the estimated number of children and adolescents having active epilepsy is 0.9M, 1.9M in ages 20-64y and 0.6M over 65 years old, with an estimated number of new cases per of respectively 130 000, 96 000 and 85 000 (2). This disease is characterized by a predisposition to generate epileptic seizures, a transient and temporary occurrence of signs/symptoms which are consequence of an abnormal neuronal activity in the brain (3).

According to the International League Against Epilepsy (ILAE) (4), since 2014, epilepsy is defined by any of the following conditions:

(1) At least two unprovoked (or reflex) seizures occurring >24 h apart;

(2) one unprovoked (or reflex) seizure and a probability of further seizures similar to the general recurrence risk (at least 60%) after two unprovoked seizures, occurring over the next 10 years;

(3) diagnosis of an epilepsy syndrome.

Epilepsy is considered to be resolved for individuals who had an age-dependent epilepsy syndrome but are now past the applicable age or those who have remained seizure-free for the last 10 years, with no seizure medicines for the last 5 years.

In April 2017, ILAE presented a revision of their previous framework (from 2010) for epilepsy classification (5). It is a multilevel classification, described in the following subsections, built so that physicians can give different levels of classification according to the available resources.

1.1.1 Seizure type

The initial point of epilepsy classification, assuming the physician already made a definitive diagnosis of an epileptic seizure, is the seizure type. Figure 1.1 shows a schematic representation of

the ILAE 2017 classification of seizure types, and describes the most common seizure types having motor or non motor onset. Seizures which include more than one type of seizure in different phases are named according to such phases. For instance, tonic-clonic seizures (defined below) have a tonic phase followed by a clonic phase.

Regarding this topic, seizures can be classified as having focal onset, generalized onset and unknown onset. Focal onset seizures start in one area in one side of the brain, while generalized onset seizures affect both sides (1).

Focal seizures may include the awareness level of the person having the seizure (i.e., if they are aware of the surrounding environment or not). Focal, generalized and unknown seizures may optionally also be subgrouped in those with motor and non-motor signs at the onset. A separate subgroup is focal to bilateral tonic-clonic seizures, which are seizures that start in one area of the brain and spread to both sides, and are different from generalized onset tonic-clonic seizures.

1.1.2 Epilepsy type

The second level is the epilepsy type, depending on the type of seizures the patient has. It is assumed that the patient has a diagnosis of epilepsy according to the 2014 definition by ILAE, presented on section 1.1 (5). Epilepsy type may be generalized, focal, combined generalized and focal or unknown.

Focal epilepsies include unifocal and multifocal disorders and seizures involving one cerebral hemisphere.

Patients who exhibit both generalized and focal seizures are characterized as having combined generalized and focal epilepsies.

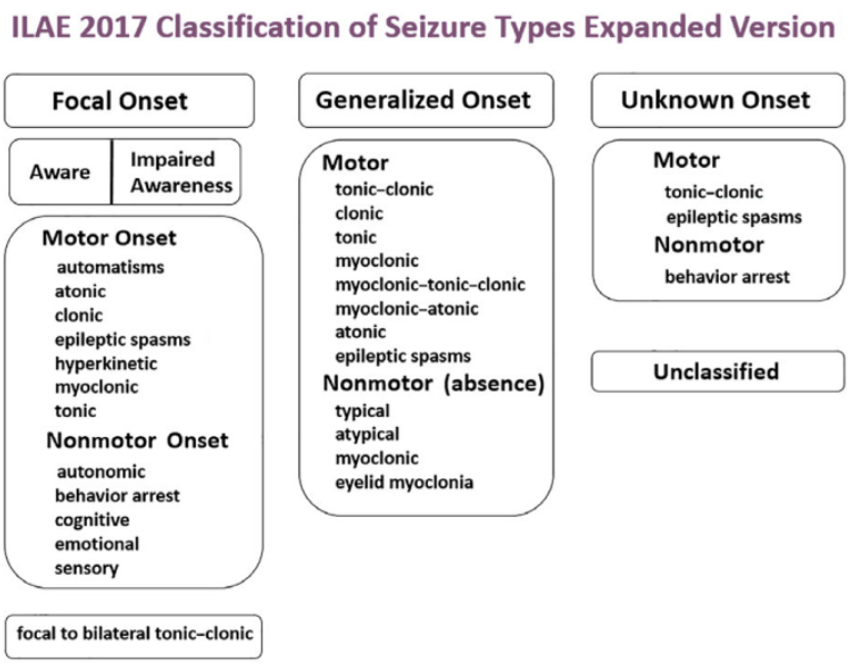
1.1.3 Epilepsy syndrome and etiology

On the third level of characterization of epilepsy (epilepsy syndrome), epilepsies are clustered by reliably identified common electrical and clinical features (5). Syndromes often contain age-dependent features (such as onset age), seizure triggers and variation according to the time of the day. Epilepsy syndromes include temporal lobe epilepsy, frontal lobe epilepsy, childhood absence epilepsy, among others.

An etiology diagnosis (cause) does not have a one-to-one correlation with the epilepsy syndrome, and its main use is to serve as guiding management for the patient's daily life. From the occurrence of the first seizure, the physician should attempt to determine the etiology of epilepsy. Etiologies can be genetic, structural, metabolic, immune, infectious or unknown.

1.2 Diagnosis in epilepsy

Current epilepsy diagnosis can be supported by a variety of diagnostic tools, such as electrophysiological recordings, functional testing, neuroimaging, and analysis of seizure semiology (study of its signs) (8).



Seizure type	Description
Atonic	Weakness or limpness of muscles
Myoclonic	Brief muscle twitching
Clonic	Sustained rhytmical jerking movements
Tonic	Increased muscle rigidity
Epileptic spasms	Repeated flexion and extension
Typical	Fixed stare, change in muscle tone and movement
Atypical	Equal to typical, but longer and may involve different symptoms
Behavior arrest	Lack of movement
Hyperkinetic/Hypermotor	Complex movements involving proximal segments of the trunk and limbs
Automatisms	Performing of unconscious actions (e.g. lipsmacking)

Figure 1.1: On top, ILAE 2017 classification of seizure types, extracted from (6). On bottom, Most common seizure types having motor or nonmotor onset (7; 1).

In some cases, it may be relevant to accurately define brain regions responsible for epilepsy. The locations and boundaries of the epileptogenic zone - the area of cortex indispensable for the generation of clinical seizures - can be defined with the help of such tools. The seizure onset zone, the area where seizures are generated, is most commonly localized by scalp, invasive electroencephalography techniques or even single photon emission computed tomography (SPECT). Typically, the seizure onset zone is inside the epileptogenic zone zone.

In general, electroencephalogram combined with video (video-EEG) is used for demonstrating the physiological manifestations of abnormal cortical activity underlying epilepsy (9). This method of measurement will be further explained in Chapter 2.

1.3 Treatment

The required treatment for epilepsy is dependent on seizure type and the diagnosed epilepsy syndrome. No cure exists for epilepsy. As such, the available treatments mainly focus on the suppression of the seizures (10).

Anti-epileptic drugs (AEDs) are the first attempted treatment in epileptic patients. Multiple trials are done to find the optimal dose and mixture for each patient (10).

If the patient does not become seizure-free with drug trials, and has a focal epilepsy, he can be subjected to resective surgery, with the removal of the epileptogenic zone. As long as it is not part of the eloquent cortex, which would lead to an impairment in function (10), its removal is possible and may lead to a significant increase in the patient's quality of life.

There are alternative, more recent treatments, such as deep brain stimulation, which directly stimulates the brain in response to a specific electrical pattern (11), mitigating seizure effects.

Epileptic patients perform long-term video electroencephalography (vEEG) in Epilepsy Monitoring Units, specialized units inside the hospital where the patient is continuously monitored in a clinical environment. Patient admission to EMUs is typically performed for differential diagnosis, seizure classification, seizure quantification, medication adjustment or presurgical evaluation (12)

1.4 Motivation

Around 60 % of the patients become seizure-free with up to three anti-seizure medication trials (13). However, about 20-30% of the patients suffering from epilepsy continue having more than one seizure per month (2), with 12% and 8% having weekly and daily seizure frequency, respectively.

Typical risks associated with seizures are injuries (due to falls or other accidents), hospitalization periods and mortality, utterly affecting the patients' mental health, which results in anxiety, depression, cognitive impairments and social exclusion, among others. The epilepsy burden not only affects the patient himself but also all the ones that live around him, namely their careers and close family (14).

It is of utmost interest to find solutions that help improve the quality of life of people with epilepsy and their caregivers.

1.4.1 Utility of seizure detection

Due to the time-limited observation period, the patient's antiepileptic medication is sometimes reduced for facilitating the occurrence of epileptic events within the observation period (15), where the seizures may not be representative of natural conditions.

In many cases, patients are asked to keep seizure diaries, as seizure monitoring is crucial for therapeutic decisions. Patient-reported seizure counts and measures derived from these reports, such as reduction in seizure frequency over a defined period, represent the primary endpoint for most clinical trials in epileptology (16).

However, it is well known that most epileptic seizure go unnoticed by the patients and their caregivers, which may affect treatment decisions. Seizure tracking is dependent on a subjective patient and family recall and may be influenced by the capacity of remembering details post-seizure, by the level of awareness during the seizure and by the ability to identify seizures (17). Accuracy of seizure detection, if based on videoEEG, can be improved by 29-fold and 7-fold, respectively, if observed by families or nurses (18). In fact, in ambulatory/home scenarios it is currently not possible to obtain valid seizure detection.

In EMUs, patient safety can be ensured by continuous observation through trained personnel. However, only around 70% of EMUs can provide this continuous patient surveillance, and only about 17% of those actually use online seizure detection and warning systems. These systems could optimize data review and make personnel assignment more efficient (19).

Moreover, by having an automatic quantification of seizures and treatment response, it should be possible to tailor the treatment for a given patient, decreasing the mortality and increasing the quality of life of patients with severe epilepsy (17).

Outside Epilepsy Monitoring Units (EMUs), seizure detection systems may help prevent severe consequences associated with seizures, such as status epilepticus (seizures with a duration over 5 minutes), injury or even death (20).

Existing devices for measuring physiological signals which have changes during seizures give enough information for providing accurate seizure quantification, and can be used for developing such systems. Table 1.1 presents a summary of the different types of seizure detection devices and their overall sensitivity.

Table 1.1: Types of seizure detection devices and their overall sensitivity. Table extracted from (17).

Seizure detector	**Overall sensitivity**
Intracranial EEG	80.5-98.8%
Scalp EEG	74-99.1%
EMG	53-100%
EDA	86-100%
ECG	33-99.8%
Accelerometer	80-100%
Video	75-100%
Cerebral oxygen saturation	57%
Mattress sensor	0-75%
Seizure alert dogs	80%
Implanted advisory system	66%

1.4.2 Utility of seizure prediction

Besides seizure detection, accurately predicting seizures a few minutes before their onset would enable patients to take precautions against injury, and could open the door for development of treatments to prevent or control the incoming seizure. For instance, neurostimulation systems can act quickly as a way of suppressing a high portion of seizures (4). This prediction should have

sufficient precision, specificity and sensitivity to minimize the interruptions on the patient's life, having minimal unpredicted seizures and false alarms.

1.4.3 Definitions

When analyzing seizure data, different time periods can be considered: interictal (between seizures), ictal (seizure periods), pre-ictal (pre-seizure periods), post-ictal (post-seizure periods) and the seizure onset. Seizure onset can either be the electrographic onset or the clinical onset. The electrographic onset is defined as the time at which the earliest signs of seizure-like activity are visible on the electroencephalogram (EEG). This onset may be followed by the clinical onset - behavioral manifestations of a seizure, as the activity spreads throughout the cortex, recognizable by an external observer or by the patient ((21), p. 140).

Seizure detection systems recognize seizures shortly before or after the actual seizure onset (determining the start of the ictal period), not providing enough time for intervention. In contrast, seizure prediction systems recognize seizures in advance of the onset zone (during the pre-ictal period), with a prediction window of up to several minutes ((21), p. 140).

Patient-specific/personalized and non patient-specific algorithms for this purpose can be distinguished. Patient-specific algorithms use the information from a large number of seizures from a given patient, and try to adjust the parameters of the algorithm for each individual. However, this requires specific training for each patient. On the other hand, non-patient specific algorithms are developed using seizure data from several patients, and as such can be applied without *a priori* knowledge regarding a patient's seizure patterns. This is useful in a clinical environment, as the same parameters can be applied for all the patients. In most cases, performance is higher for patient-specific algorithms (19), at the disadvantage of requiring previous data acquisition.

1.5 Economic impact of solution

The economic impact of having a proper solution for seizure detection and prediction is estimated in this section.

Previous surveys have shown that in Europe and Israel, there is at least 150 EMUs (22). A review study including EMUs from 25 countries showed that the average number of beds per EMU was 6.8 (standard deviation - SD - 11.1), reporting an average length of stay of 4.8 days (SD 1.8) (23).

European EMUs can monitor children, adults, or both. 13% of them monitor only adults, with less than 50 admissions per year. 13% monitor only children, with 50-150 admissions per year. 74% monitor both, with 28% having 150-250 admissions per year and 19% with over 250 admissions (22).

In a review of the cost of brain disorders in Europe, the total annual cost per patient with epilepsy ranged from 695€ to 11.654€ (24). The total costs are estimated to be 13.9 billion per year (25). The variability in cost estimation is due to study design, source population heterogeneity

of and the difficulty in separating the costs related with underlying epilepsy causes and comorbidities and its intrinsic costs. The annual costs per patient were found to decrease on the subsequent years of follow-up after seizure diagnosis, and are higher in patients with drug-resistant epilepsy, surgical candidates, patients with disability, different seizure types, and status epilepticus. The more expensive clinical conditions are mostly represented by epilepsies with uncontrolled seizures requiring treatment changes and more frequent use of newer antiepileptic drugs. With objective seizure quantification, these costs could be reduced by quickly assessing the efficacy of a new drug on ambulatory.

Regarding the costs of injuries, in Scotland, a study surveyed admissions of patients with trauma caused by seizures from a population of over 400 000 (26). 12 patients had 13 admissions to the hospital, with 14 injuries caused by seizures for around 3.5 years. The cost was estimated to be at around 4400€ per traumatic incident. Seven of those patients may have had potentially preventable seizures, with more closer control of their epilepsy and medication adherence.

Furthermore, status epilepticus has an estimated cost of €8347 per admission, in Germany, projected at an yearly value of €83 million (27). Patients with status epilepticus are hospitalized for significantly longer periods (average length of 13 days), compared to other epileptic disorders and acute illness.

Table 1.2 contains the average drug and hospital stay costs per European country, differentiated by type of cost.

Table 1.2: Average costs of hospital stay and antiepileptic drugs.

Country	Sample Size	Median cost per year, per patient
Russia (24)	738 (adults)	Drugs: €643 (range €0–9960) Hospital costs: €647-€950
Northern Italy (28)	525 (children+adults)	Drugs: €827 (range €171-1815) Hospital costs: €552 (range €412-3950)
France (29)	405 (adults)	Drugs: €2298 (range €133-8145) Hospital costs: €991 (range € 0-22 230)
Germany (30)	366 (adults)	Drugs: €1017 (range €23-8334) Hospital costs: €510 (range €0-13 091)

It can be concluded that seizure detection and prediction devices can help mitigate epilepsy related costs, and should be very impactful in doing so. Summarizing, injury can be prevented or its severity can be reduced, therefore reducing costs related with hospital stays. Drug-related costs can be reduced by having a better assessment of which drugs are effective on ambulatory. Lastly, clinical trials can be improved by a more precise estimation of reduction in seizure frequency.

1.6 Objectives

The present objectives are based on the line of research "MovEpil" of the BRAINLab R&D group from INESC-TEC where this dissertation was developed.

The main goal of this dissertation is to create and study the impact of a system for multimodal epileptic seizure detection and prediction using electrocardiogram and eletroencephalogram data acquired in Epilepsy Monitoring Units. In the future, the goal is to transfer that solution to a wearable multimodal device, with the best sensors chosen according to each patient. In this dissertation, the potential of having this solution translated into a portable system will also be assessed by implementing the developed architecture in an embedded AI computing single board system (NVIDIA Jetson TX1).

1.7 Main Contributions

The main contributions from this dissertation are the following:

- Presented results for the first combination of deep learning and multimodal data fusion for seizure detection and prediction, and for the first multimodal seizure prediction system;
- Showed an innovative approach for data augmentation, applicable for this domain in EEG and ECG.
- Got accepted for oral presentation on MEDICON International IEEE-sponsored conference (2019), with a paper titled *Multimodal approach for epileptic seizure detection in Epilepsy Monitoring Units* (in appendix B). The paper was reviewed with 2 and 3 by two reviewers, on a scale ranging from -3 to 3.
- Assisted INESC-TEC's BRAINLab team in writing an H2020 proposal for a Research and Innovation Action (ECSEL 7 May 2019 call), for the project **WePA-EPil** - Towards a Personalized intelligent Wearable & Portable Ambulatory system for better EPilepsy seizures detection and management, with a total budget of around 600k € (300k€ for INESC TEC). The other consortium partners were Dengun Lda, epihunter, and Ludwig Maximillian University of Munich. My responsibilities were partially writing some workpackages, researching the state of the art, impact of the solution, video-chatting with consortium partners and proposing changes in the manuscript;
- Assisted the same team in writing an H2020 proposal for an Innovation Action (FTI 23 May 2019 call), for the project **MIST -** Multimodal Intelligent wearable System for epileptic Seizure Tracking, with a total global budget of 3M € (around 750k € for INESC TEC). The other consortium partners were Dengun Lda, epihunter, Ludwig Maximillian University of Munich and Neuro Event Labs Oy. My responsibilities were similar to the described above.

- Participated in a workshop organized by Fraunhofer's MDevNet, "Prospeção e Identificação de Tecnologias com Elevado Potencial de Transferência para o Mercado" with the NeuroKinect technology from BRAINLab (3D VideoEEG including quantification of seizures). NeuroKinect was partially used for the previous recording of seizure data in this dissertation (appendix C)

- Proposed a protocol (appendix D) to be used in the future in the Epilepsy Monitoring Unit, for recording biosignal activity when the patient is performing specific movements of interest. This protocol was planned as a way to reduce the number of false positives in an algorithm, since artifacts derived from normal movements can be more easily recognized.

1.8 Structure

The remainder of this work is structured as follows: Chapter 2 presents an overview of the main physiological changes during seizures, and the ways of measuring them. Chapter 3 reviews the methods and techniques most commonly used in the literature regarding seizure detection and prediction. Chapter 4 shows the state of the art approaches for each step of the classification pipeline. Chapter 5 presents the results for a multimodal seizure detection system, and proposes some solutions for helping physicians understand the system. In Chapter 6, the developed algorithm is implemented on NVIDIA Jetson TX1 and validated in real time. In Chapter 7, a multiclass classifier, for seizure detection and prediction, is trained on the available data, for obtaining preliminary results on seizure prediction. Finally, the dissertation's conclusion and future work are presented on Chapter 8.

Chapter 2

Modalities for seizure detection and prediction

This chapter aims to give an overview on the different modalities for seizure detection and prediction. In the following sections, the most commonly used sensors in a clinical environment for seizure analysis (EEG, ECG, EMG, EDA and movement) will be described, along with the corresponding physiological changes during epileptic seizures, and advantages of using each sensor.

2.1 Electroencephalogram

Surface electroencephalography (EEG) is an electrophysiological monitoring method used for recording the electrical activity of the brain (electroencephalogram). Combined with video, EEG is the golden standard for seizure diagnosis (9) in Epilepsy Monitoring Units.

Surface EEG signals can be broken down into the following bands: delta (0.1-3.5 Hz), theta (4-7.5 Hz), alpha (8-13Hz), beta (14-30 Hz) and gamma (above 30 Hz) (31). The signal amplitude ranges from 5 to 300 μV, and its frequency may achieve values up to 150Hz. The occurrence of epileptic seizures leads to alterations in these normal frequencies and amplitudes (9).

As a way to ensure cross-study reproducibility, the 10/20 or 10/10 system are commonly used for standardizing electrode placement in EEG recordings (32). Figure 2.1 shows a single plane projection of the head with the positions of the electrodes, for both systems.

On the 10-20 system, composed of 21 electrodes, each electrode has a code which identifies the lobe and the hemisphere. Letters F, T, P and O represent frontal, temporal, parietal and occipital lobe, respectively. C are the electrodes in the central position. Odd numbers indicate electrodes in the left hemisphere, while even numbers correspond to those in the right hemisphere. Finally, the numbers 10 and 20 correspond to the percentages of distances between adjacent electrodes measured according to standard landmarks of the skull, in a given plane (32).

The 10/10 system is an extension to the 10/20 system, which can capture with higher detail the propagation of electrical activity in the scalp during seizures (33). In this system's nomenclature, two letters are combined to indicate a mid point between two regions (e.g. the electrodes between

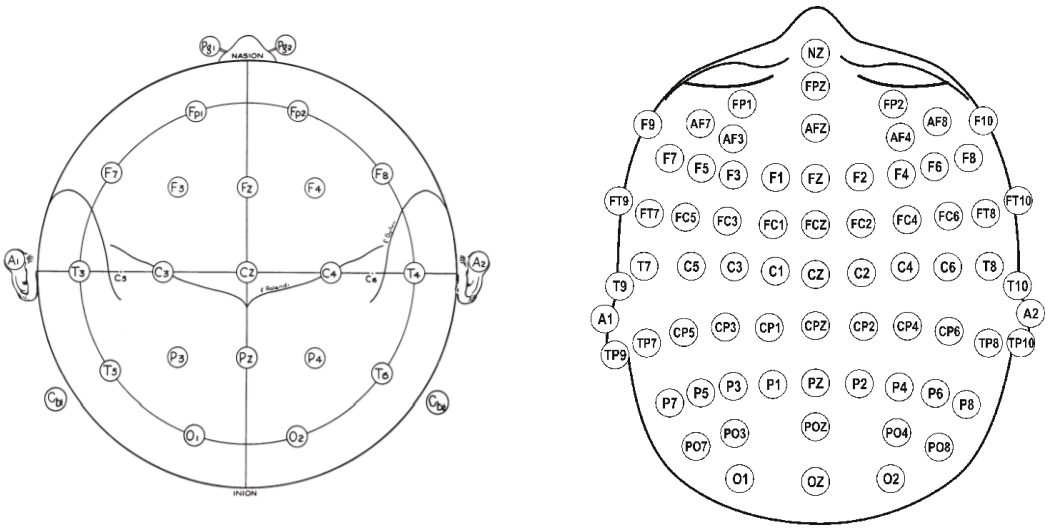


Figure 2.1: Single plane projection of the head, showing the locations of the electrodes according to the 10/20 (left) and 10/10 (right) system. Figures extracted from (32; 33).

the P contour and the C contour are named CP). Moreover, between the Frontal Pole (Fp) locations and the frontal, the new electrode positions are prefixed with "A", as they are placed from the anterior to posterior position.

Epileptic brain activity can be recorded during the ictal or interictal period. For generalized seizures, spike wave patterns consisting of polyspikes, spike runs or rhythmic beta activity are typically seen. In focal epileptic seizures, low amplitude, high frequencies, attenuation of ongoing EEG activity or sharply contoured theta waves are commonly observed ((21), p. 55).

While performing EEG-based seizure analysis, a broad range of seizures in patients with different seizure onset areas and different epilepsy syndromes can be detected. In a clinical setting, multiple EEG channels are commonly used, as this leads to an increase in the sensitivity. However, the computational cost for seizure analysis increases with the number of channels, and the patient comfort decreases, especially in an outpatient environment (19).

Regarding its limitations, surface EEG has a very good temporal resolution but a bad spatial resolution. As such, recordings may miss out on underlying epileptic patterns (34). In terms of usability, EEG electrodes can be uncomfortable if a good signal quality is required, such as this case, and are not aesthetically pleasing.

2.2 Intracranial Electroencephalogram

Intracranial EEG (iEEG) is used in cases where non-invasive methods result in poor seizure localization. For instance, in pre-resection surgery evaluation, it is required that the seizure onset area is correctly localized so that the physicians can know whether or not that area can be safely removed (35). iEEG's amplitude ranges from 10-5000μV, with frequencies from 0 to 150Hz ((36), p. 954).

The major benefits of iEEG compared to EEG are it being less prone to contamination with artifacts and having greater proximity to sites of seizure generation (21). The changes during

seizures are similar compared to surface EEG. As such, iEEG is typically used for finding predictive features of seizures.

Contrary to EEG, the electrode placement and the number of electrodes in iEEG is not standardized, and is patient-dependent (37).

iEEG has the major disadvantage of requiring surgery for implantation. Moreover, current intracranial EEG recordings in hospitals last 7-10 days (limited by infection risk, hospital costs and patient compliance), and the implantation of iEEG electrodes leads to alterations on seizure dynamics which may not occur after a few months of use (38).

2.3 Electromyogram

The electromyogram (EMG) is the electric potential generated by skeletal muscle cells when they are activated. Surface EMG signals have an amplitude between 0.1–5 mV and a frequency range of 2–500 Hz ((36), p. 954) and are measured using surface electrodes. An example of an EMG signal during a generalized tonic clonic seizure is shown in figure 2.2.

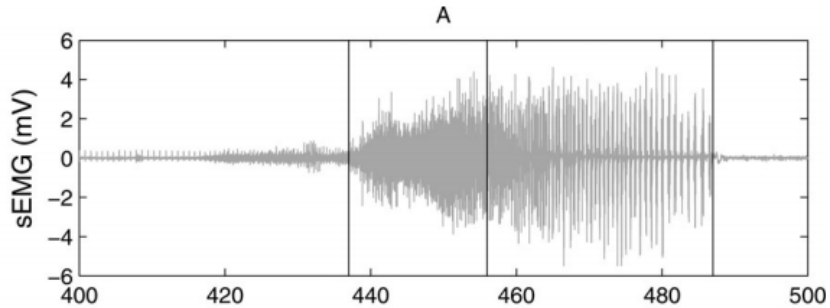


Figure 2.2: Seizure activity in an EMG during a generalized tonic clonic seizure. Figure extracted from (39).

As most seizure types have a motor component, surface EMG can be a viable option for seizure detection. The anterior tibialis, deltoid, triceps and brachial biceps muscles have been used as recording sites for seizure detection (40; 41; 39; 42). The placement of the electrodes may be patient dependent, as depending on the seizure onset zone, different muscles are triggered during the motor seizure.

Motor seizures have a high activity in the frequency band above 100–150 Hz, when compared to non-epileptic seizures (39). A quantitative analysis of tonic (TS) and tonic-clonic (TCS) seizures showed differences when compared to physiological muscle activation, expressed as the maximum voluntary contraction (MVC) during simulated seizures by healthy volunteers. The tonic phase in TS was even different from the tonic phase of TCS: EMG signals in TS had higher frequency compared to tonic-clonic and MVC, and TCS had a higher amplitude compared to TS and MVC (43).

While EMG electrodes are more comfortable for daily use compared to EEG electrodes, they are prone to more false positives, as normal motion may be confused for epileptic activity.

2.4 Electrocardiogram

ECG (also known as EKG) is a biosignal originated by the electrical changes on the skin that arise from the heart muscle's depolarization and repolarization during a heartbeat. ECG signals have an amplitude between 1-10 mV and a frequency range of 0.05–100 Hz ((36), p. 954). An example of an ECG segment is shown in figure 2.3.

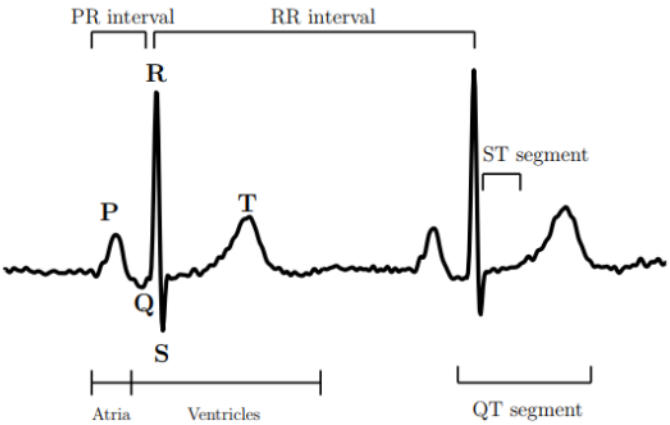


Figure 2.3: Example of an ECG segment, with different key points signaled. Figure extracted from (44).

Typically, in clinical practice, 12-lead ECG sensor placement are used. However, in EMUs only one lead is chosen, to diminish information overload, as it contains most of the necessary information for characterizing seizures (44).

In 82% of the patients with epilepsy, seizures are accompanied by ictal sinus tachycardia (45), due to the activation of the central autonomic network caused by epileptic discharges. This occurs especially in patients with generalized tonic-clonic seizures and focal impaired awareness seizures, originating from the temporal lobe (46). The ictal heart rate changes are also stronger in these patients. Ictal tachycardia can also be caused by the motor manifestations during seizures, but to a much lesser extent (47). Compared to heart rate changes caused by physical exercise or nocturnal arousal, changes during epileptic seizures are faster and more pronounced (19; 21).

The definition of ictal tachycardia is variable in the literature. While the most frequent definition is having a heart rate over 100 beats per minute (bpm), other works define it as over 120 beats per minute (bpm), or even having 10bpm above the baseline (45).

Figure 2.4 shows the occurrence of alterations on the ECG on par with the build-up of EEG activity.

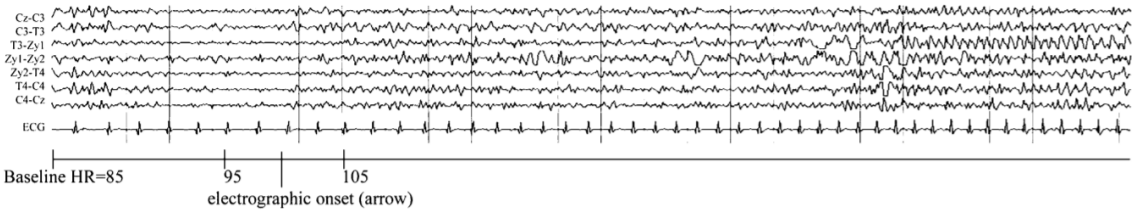


Figure 2.4: Visible heart rate changes during a seizure. Figure extracted from (48).

In some cases, heart rate changes can occur even before the earliest electrographic or clinical change (48; 49), making them suitable for real-time usage as well as early seizure detection/prediction. In fact, the median onset of pre-ictal heart rate increases reported in a study was 10.7 s prior to the seizure onset (50).

Ictal heart rate changes typically follow a specific pattern, as represented in Figure 2.5. First, the heart rate increase linearly, stabilizes at a certain peak value and finally decrease exponentially (51). However, there is a high variation within each patient, in both shape and amplitude.

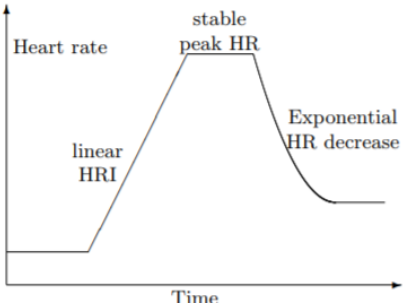


Figure 2.5: Illustration of ictal heartrate changes, extracted from (51).

Besides heart rate changes, around 40% of patients have ictal ECG abnormalities such as atrial fibrillation, asystoles, changes in QRS morphology or even T-wave inversion (52).

ECG is highly robust and less susceptible to artifacts, compared to EEG. In an ambulatory setting, long-term recordings can be easily obtained, with minimally invasive approaches, but causing some discomfort. ECG can be recorded with only one channel, being much simpler to analyze and to process than EEG. However, heart rate changes occur in many everyday activities, and are not characteristic of all seizure types, leading to some seizures being undetected and to a high susceptibility for false positives, if the algorithm is not correctly tuned for each patient (17).

2.5 Electrodermal activity (EDA)

Sweat changes occur during seizures, as a reflection of the activity of the sympathetic branch of the autonomic nervous system, altering skin condutance (EDA).

The disadvantages of using the electrodermal activity include that its recording is susceptible to motion and pressure artifacts, and can be obtrusive or uncomfortable (17).

2.6 Movement sensors

In clinical practice, movement is the first of all non-EEG seizure detection to be broadly used, as it is very intuitive (53).

For movement recording, several sensors are available. Contact sensors (such as accelerometers) have an increased sensitivity compared to non-contact sensors (such as video) but are more intrusive to the patient (53).

Accelerometers can only be used in a select portion of seizures that have well defined motor activity, and may have a lot of false positives when used in the daily environment as quick movements may be similar to seizures. Current video detection systems are limited by being applicable only in area that is covered by the camera and by the inability of capturing seizure patterns of events that occur under view occlusion, such as blankets (54).

2.7 Multimodal detectors

Recently, interest has risen in combining several of the above described sensors for observing physiological changes in a higher percentage of seizure types. In fact, many EMUs, such as Ludwig Maximillian University of Munich, already measure ECG synchronized with videoEEG.

Table 2.1 shows the occurrence of several physiological (non-EEG) alterations on the most common seizure types, during the ictal period.

Table 2.1: Main findings of the occurrence (+) or not (-) of movement, sweating or heart rate (HR)/ECG changes for the main seizure types. Table adapted from (17).

Seizure type	**Movement**	**Sweating**	**HR/ECG changes**
Atonic	±	*	*
Autonomic	-	+	+
Clonic	+	+	+
Myoclonic	+	*	-
Epileptic spasm	+	*	*
Focal onset impaired awareness seizures	±	+	+
Generalized tonic-clonic	+	+	+
Hypermotor	+	+	+
Tonic	+	+	+

* No data on this particular issue

There is not a single non-EEG physiological signal that can detect all seizures. As such, the best solution for seizure detection and prediction should be in combining several physiological signals depending on a patient's seizure type, which may include EEG.

2.8 Required needs for a wearable device

In the literature, the most investigated possibility for a marketable device is to have an automated seizure warning system. This system would, for instance, send the GPS location of the epileptic

patient to the caregivers, reducing the chance of severe consequences. To be clinically accepted, the devices should have high sensitivity, low false alarm rate and low seizure detection delay.

An improvement in the algorithm's sensitivity leads to a decrease in its false alarm rate - as such, a balance according to seizure type is required. Seizures with violent motor manifestations (such as tonic-clonic seizures) are more easy to detect, and patients accept a device having a minimum sensitivity of 80% (55). On the other hand, for clinically less dangerous seizures (such as focal seizures with automatisms) a sensitivity from 70 to 80% is acceptable.

Detection delay is dependent on the time it would take for the caregivers to go from their current position to the patients'. If the caregiver is not in the same house, an extra delay in exchange for higher performance is not critical (56). In that scenario, the effect of false alarms is also higher, as would lead to an unnecessary trip.

More practical concerns, such as the device's autonomy, also need to be addressed. High computational power and high autonomy also require a high device size (57).

Finally, premarketing surveys indicated that patients generally appreciate the development of personal devices for seizure detection and prediction. Most of them would prefer technologies that are already used by healthy people, such as smartwatches and fitness trackers (in the form of wristbands or chest straps). Invisible patch sensors for ECG or sEMG would also be acceptable. Patients have little interest in using implanted electrodes for seizure detection (16).

Chapter 3

Methods for seizure detection and prediction

The literature of seizure detection and prediction shows a high variety of different methods employed in different circumstances. In all of them, the signals are acquired, filtered and digitized with appropriate equipment, for a variety of patients and seizures. The most common methodologies are described in this chapter, showing the possible pipelines and different steps that can be taken to solve this issue. In the next chapter, the most relevant papers using such methods specifically for epilepsy will be presented.

3.1 Machine Learning and Model Evaluation

Machine Learning (ML) is a field of artificial intelligence dedicated to pattern recognition by learning from given data. As such, the goal is to learn a function that can predict a variable y from features extracted from data, X. Seizure detection and prediction is a classification problem in which the output y is discrete, whose range depends on the number of classes we intend to classify the signal into (58). For instance, classifying the signal into 'ictal/non-ictal' state would be a binary classification problem.

In the classical machine-learning pipeline, appropriate features are extracted from the signals, according to their known characteristics during the period of interest, as a way of increasing the contrast between ictal and non-ictal periods (in the case of seizure detection) or pre-ictal and interictal periods (for seizure prediction).

The main issue with classical machine learning is that it is very target-directed, and there is no way of knowing *a priori* which feature extraction algorithm is best for a given problem. A certain feature is considered to be good based on experimental approaches. As such, it requires a long time to optimize the best features for a given problem (59).

3.2 Deep Learning

In the past few years, deep learning techniques have achieved very good results in challenges that conventional machine learning techniques could not solve, due to an increase in computing power and the elimination of the need for feature extraction and selection. Besides, deep learning algorithms extract optimal features intrinsically, considering all the available evidence, leading to a more natural and unbiased representation (59).

The nomenclature *deep* is derived from the numerous number of layers in the structure of an Artificial Neural Network (ANN).

3.2.1 Convolutional Neural Networks

A feedforward neural network, also called Multilayer Perceptron (MLP), is built by basic blocks called perceptrons/neurons, which are linear classifiers. It is a set of connected input/output units where each connection has a weight assigned to it, which is adjusted during training.

Convolutional Neural Networks (CNNs) are 3D volumes of neurons that perform a series of convolutions, sharing weights.

A simple CNN is typically built according to the following architecture (60):

- INPUT: holds the raw values of the signal;
- CONV layer: consists of a series of filters with learnable weights that can be stacked spatially. These filters are convolved with the input sample, acting as feature extractors whose weights are continuously updated.
- ACTIVATION layer: Typically used after a CONV layer. It applies an elementwise activation function.
- POOL layer: performs a downsampling, reducing the input sample size, and retaining the most important information. The most commonly used operation is max pooling, which maintains the maximum value of a given neighbourhood.
- FC (fully-connected) layer: each neuron in the previous layer is connected to all the neurons in the current layer. The number of classes of the classification problem is the total number of fully-connected neurons in the final layer.
- Batch Normalization layer: consists on normalizing the input of each layer by adjusting and scaling the activations. It does so by subtracting the batch mean and dividing it by the standard deviation. BN reduces the amount of shifting of the hidden unit values (covariance shift), while also adding some regularizing effect to the classification.

3.2.2 Long short-term memory (LSTM) networks

Other common types of structures for Deep Learning are Recurrent Neural Networks (RNN) (61). RNNs apply a recursive approach where the output is dependent on the previous computation,

as represented in figure 3.1. In practice, RNNs do not have the capability of learning long-term dependencies, something that may be of use for learning patterns of the pre-ictal state.

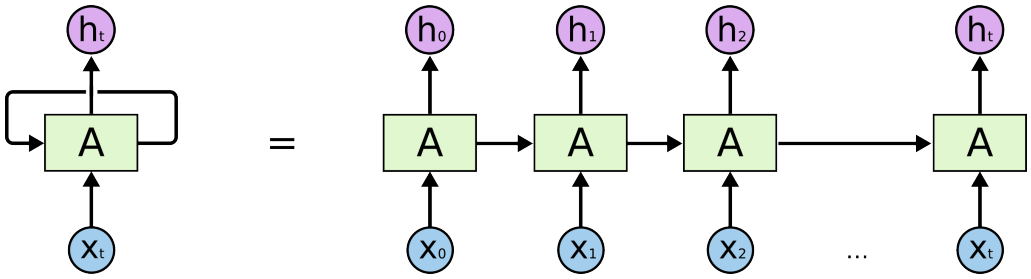


Figure 3.1: Graphical representation a Recurrent Neural Network. Figure extracted from (62).

To remember information for a long period of time, LSTMs, a subtype of RNNs, were developed. LSTM layers have an oriented connection throughout all its units, representing time instances. These layers contain a memory block with three gates - input, output and forget - which decide what information should be stored/updated, used or discarded, respectively (63). All gates process their input based on the memory from the previous unit and the current state of memory.

3.2.3 Stacked Sparse Denoising Autoencoders

An autoencoder is a neural network that learns the necessary weights for setting the target values to be equal to the inputs. By doing so, the algorithm's middle layers end up discovering correlations between features (64), working a a feature selector.

Denoising autoencoders, an extension of the basic autoencoder, consist on autoencoders where a small amount of noise is added to the input, and the networks attempts to obtain the original, denoised image. This is advantageous compared to the previous approach, as standard autoencoders may learn the identity function with no dimensionality reduction (65).

In sparse autoencoders, the network is forced to selectively activate regions depending on the input data. The network's capacity of memorizing the input data is limited, without limiting the networks capability to extract features from the data (by reducing the number of nodes).

Sparse and denoising autoencoders can be combined and stacked, building a Stacked Sparse Denoising Autoencoder (SSDA) network (64), which has the advantage of capturing a hierarchical grouping of the input.

3.2.4 Activation functions

Activation functions, used for adding non-linearity to neural networks, are applied to vectors' output from neural networks before computing the loss. The most commonly used activation functions are sigmoid, softmax and ReLu.

Sigmoid function (Eq. 3.1) squashes the output vector, s_i, in the range [0, 1]. It is applied independently to each element.

$$f(s_i) = \frac{1}{1+e^{s_i}} \tag{3.1}$$

In softmax (Eq. 3.2), all the scores add up to 1. Elements can be interpreted as class probabilities, as they represent a class.

$$f(s_i) = \frac{e^{s_i}}{\sum_j^C e^{s_j}}, \tag{3.2}$$

where s_j are the scores inferred by the network for each class C.

ReLu (Rectified Linear Unit - Eq. 3.3) is a linear activation function, most commonly used in CNNs, with fast convergence and easy computation.

$$f(s_i) = maximum(0, s_i) \tag{3.3}$$

3.2.5 Loss Functions

The cross-entropy loss, defined by Eq. 3.4, is the most common loss for classification problems.

$$CE = -\sum_i^C t_i log(s_i), \tag{3.4}$$

where t_i and s_i are the ground truth and the classifier score for each class i in C.

Typically, an activation function (Sigmoid/Softmax) is applied to the scores before calculating the cross-entropy loss. For binary classification, sigmoid is typically applied (and the loss function is named binary cross-entropy), while in multiclass classification, softmax activation is commonly used (obtaining the categorical cross-entropy loss function).

Based on the building blocks presented (CNNs and LSTMs), very complex architectures for classification can be developed.

3.3 Dimensionality reduction

In the attempt to capture the wide variety of information in both ictal and pre-ictal periods, the number of extracted features can easily become too high for real-time seizure detection and prediction. While a high number of features may help discriminate more complex patterns, there is a greater potential for overfitting, especially in small datasets. Besides, irrelevant features can also blur the boundaries between classes (66).

Dimensionality reduction techniques can determine which features are more discriminative for the data available, and even for each patient, building patient-specific algorithms.

The following sections describe the most common algorithms for dimensionality reduction.

3.3.1 Principal Component Analysis (PCA)

Principal Component Analysis (67) is a feature projection method which uses an orthogonal transformation for transforming a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables (principal components). The first principal component has the highest variance in data, and each succeeding component has the highest variable possible while being orthogonal to the preceding components. During this method, it is possible to choose the amount of variance (typically, 95% or 99%) or the amount of features which want to be maintained.

3.3.2 Sequential/Backwards Feature Selection

Sequential and Backward feature selections are iterative algorithms which may be used with either filter or wrapper methods (68). The Sequential Feature Selection algorithm starts with an empty feature set, adding one feature that gives the highest value for a certain objective function. On the next iteration, features are added individually to the current subset, and each feature is included if it gives the maximum value for the objective function. The process is repeated until the required number of features is added. A Sequential Backward Selection algorithm is similar but starting at a complete feature set, and removing the features that give the lowest decrease on the objective function's performance.

3.3.3 ReliefF

ReliefF (69) is a feature selection algorithm, sensitive to feature dependencies. This algorithm starts by assigning a weight of zero to each feature, and then works iteratively, as follows: a random data point is sampled, and a sample of the closest k data points from the same class and from other classes are found. Then, the feature weights are adjusted as a way to enhance the features that seem more discriminative. The algorithm penalizes the features that give distant values to neighbours of the same class, and rewards those that give different values to neighbours of different classes. The output is a vector of feature weights, which can be used for feature ranking.

3.4 Multimodal data fusion

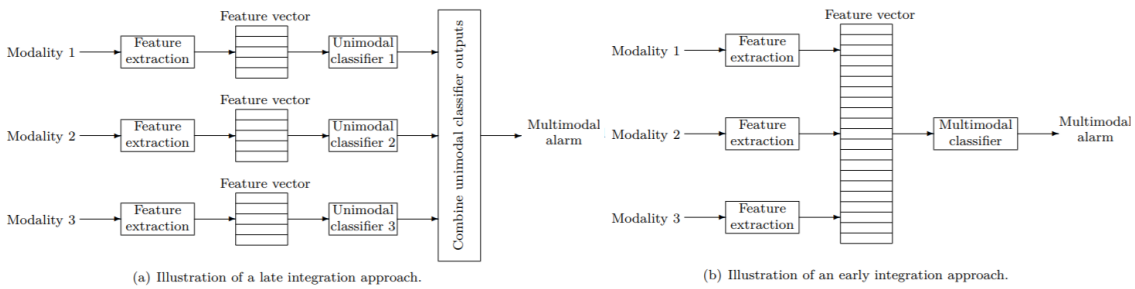
When using more than one type of signal, the features extracted from each channel or the classification output must be fused (52).

A late integration system is a system where a decision is done per modality/sensor. Each unimodal decision is combined (typically, using *AND* or *OR* logical operators) and a final decision is given. This is particularly useful when one of the signals, such as EEG, shows artifacts that resemble seizures, which would increase the number of false alarms if unimodal detection was being used.

Combined with late integration, *override* can be used. The *override* feature overrides the previous approach when the detector observes consecutive global decisions, which is important when manifestations from other biosignals (such as ECG, where ictal tachycardia is seen) are slower and occur only a few seconds after onset.

In contrast, in early integration systems, the features are extracted from each modality and combined in one general feature vector. These approaches do not use any knowledge of the delay of physiological changes seen in each biosignal. As such, late integration is most typically seen in the literature.

A schematic representation of both approaches is shown in Figure 3.2.



(a) Illustration of a late integration approach.

(b) Illustration of an early integration approach.

Figure 3.2: Early and late integration for multimodal approaches. Figure extracted from (44).

3.5 Evaluation metrics

For seizure classification, depending on the actual state of the epoch being classified, one out of four outcomes are possible: a true-seizure period is missed (false negative - FN), correctly detected (true positive - TP), wrongly flagged (false positive - FP) or rejected if it is not a seizure period (true negative - TN). This is applicable in both seizure detection and prediction ((21), p. 154).

The following subsections describe common performance metrics used for seizure detection and prediction.

3.5.1 Detection latency (DL)

The detection latency metric (DL) is the time interval from seizure onset to actual detection, and as such, is directly related to detection speed. Depending on how much of the signal we need to see before making the optimal decision, the detection latency varies. For larger windows, we may have some more smoothing of the features (if we apply operations like the mean), but we may miss brief bursts of characteristic activity of the ictal period.

When applied to seizure prediction, delay is negative (and typically mentioned as prediction time).

3.5.2 Sensitivity (Sens)

Sensitivity (Eq. 3.5) expresses the likelihood that a true seizure is detected by the seizure detector, or that a true preictal period is detected by the seizure predictor.

$$Sens = \frac{TP}{TP+FN} \tag{3.5}$$

3.5.3 Specificity (Spec)

A detector may be extremely sensitive, but generate a lot of false positives. Specificity (Eq. 3.6) is a metric which relates the selectiveness of the seizure classifier to seizure patterns. In the case of seizure classification, it tells how unlikely the classifier is of mistaking interictal activity for seizures (in seizure detection) or interictal/ictal activity for the pre-ictal period (in seizure prediction).

$$Spec = \frac{TN}{FP+TN} \tag{3.6}$$

3.5.4 False Positive Rate (FPR)

The False Positive Rate (FPR) is defined as the number of false positives occurring per unit time (typically, by hour or by 24 hours). When analysing the FPR of a classifier, the true seizure rate should also be considered - for instance, a FPR of 0.1/h may be unacceptable if the patient only has a seizure per week.

3.5.5 Positive Prediction Value (PPV)

Since specificity is defined with respect to the number of true negatives, which may be difficult to assess because EEG does not monitor electrical activity in all possible locations, positive prediction value (Eq. 3.7) gives information about the fraction of detection that are true positive, without telling us how many seizures are likely to be missed.

$$PPV = \frac{TN}{FP+TP} \tag{3.7}$$

3.5.6 F1-score

The F1-score (Eq. 3.8) is a metric which is commonly calculated in classification problems, and is a measure of a test's accuracy.

$$F1 = \frac{2*TP}{2*TP+FN+FP} \tag{3.8}$$

3.5.7 Area Under Curve (AUC)

The receiver operating characteristic (ROC) curve is a graphical analysis of the trade-off between sensitivity and specificity of a classifier, obtained by varying a parameter such as the classifier's

probability threshold, which is 0.5 by default. The area under the ROC curve gives us a measure of this trade-off.

3.5.8 Seizure Prediction Horizon and Seizure Occurrence Period

The Seizure Occurrence Period (SOP) is defined as the interval where the seizure is expected to occur. The period between the beginning of the prediction alarm and seizure occurrence period (SOP) is called Seizure Prediction Horizon (70) (SPH). The seizure onset must be within the SOP and after the SPH in order to be correctly predicted.

Chapter 4

State-of-the-art

On PubMed, a search for *(((seizure detection) OR seizure prediction) OR (EEG seizure detection and prediction)) OR (ECG seizure detection and prediction)* returns a total of 4177 results. From 2008 to 2018 (except in 2017), there has been an increase from 116 to 328 papers per year with these keywords.

From the initial 4177 papers, the ones published starting at 2014 were selected, in English (1625 papers). From those, 234 are review papers. The most recent review papers and with more citations, for each of the modalities (EEG, ECG, multimodal) were then picked. From each review article read, based on the cited papers' abstracts, the most different papers, with best results or a new approach were chosen. 12 papers using EEG or iEEG were selected. Two more were added from a later search due to their innovative, "out of the box" approach based on deep learning, although they were not peer reviewed. From ECG and multimodal approaches, 9 and 4 papers were selected, respectively.

The following sections present a summary of those chosen key papers, for providing a review on state of the art on seizure detection and prediction using ECG and EEG.

4.1 Available datasets

Datasets for studying seizure prediction and detection should contain a lot of seizures, and a sufficient time period between consecutive seizures, so they can be considered independent events. If they are clustered, it becomes hard to separate the presumed pre-ictal from the postictal period on EEG data, since their exact duration is unknown (71).

A search for datasets used for seizure detection and prediction, whose result is presented next, was performed by combining the results from the above papers and Kaggle competitions related with epilepsy.

4.1.1 Bonn University

Bonn University dataset (72) is widely used for epileptic seizure detection. It contains five different sets (A-E). A and B include 23.6 seconds long, 100 single-channel EEG signals, from 5 healthy

participants, using the 10-20 system for electrode placement. In set A, participants have their eyes open, while in B they are closed. Sets C and D are iEEG data taken from five epileptic patients, in interictal periods. Set E includes EEG data from five patients in the ictal period.

4.1.2 UPenn and Mayo Clinic

In 2014, a Kaggle [1] competition sponsored by National Institutes of Health's National Institute of Neurological Disorders and Stroke (NINDS) and the American Epilepsy Society (AES) was hosted to develop the best personalized seizure detection algorithms (73). This competition included four iEEG datasets recorded from dogs (348 seizures recorded over 1,500 days) with naturally occurring epilepsy, using the NeuroVista Seizure Advisory System, and eight iEEG datasets from human patients (942 seizures recorded over more than 500 days) with drug-resistant epilepsy, being followed at Mayo Clinic Rochester.

In these datasets, ictal and interictal data segments were labeled. The same datasets, but labeled on preictal and interictal segments, originated a different challenge for seizure prediction, in the same year.

4.1.3 Epilepsy Center of the University Hospital of Freiburg

The iEEG database from the Epilepsy Center of the University Hospital of Freiburg (74) contains recordings of 21 patients suffering from medically intractable focal epilepsy. This database is now discontinued, being part of a larger project named The European Epilepsy Database (75), which contains annotated EEG datasets from more than 250 patients with epilepsy. However, only part of this dataset (30 patients) is available, after paying a fee of $3000.

4.1.4 Melbourne University AES/MathWorks/NIH Seizure Prediction

In a 2016 Kaggle competition, participants were challenged to distinguish between ten minute long data clips covering an hour prior to a seizure, and ten minute iEEG clips of interictal activity. As such, the goal was to develop a seizure prediction system based on long-time acquisitions. 1,139 seizures were recorded over 1,326 days, from three humans, from the world-first clinical trial of the implantable NeuroVista Seizure Advisory System (76).

As seizures are known to cluster/occur in groups, patients who typically have seizure clusters receive little benefit from forecasting follow-on seizures. As such, in this dataset, lead seizures (occurring four hours or more after another seizure) were included in the training and testing data sets. Interictal data segments were restricted to be at least four hours before or after any seizure, to avoid contamination between interictal, pre-ictal and postictal segments. Pre-ictal data is available for an hour before the seizure.

Ictal data from the NeuroVista trial, from 12 subjects, is also available (76), but with no pre-ictal correspondence, making this unsuitable for seizure detection, but useful for understanding seizure patterns.

[1] Kaggle - www.kaggle.com.

4.1.5 The Temple University Hospital EEG Data Corpus

The Temple University Hospital (TUH-EEG Corpus), is an ongoing data collection effort that has recently released 14 years of clinical EEG data collected at Temple University Hospital. It contains 16,986 sessions from 10,874 unique subjects, comprising a grand total of 29.1 years of EEG data (2016 version) (77). There is a subset for seizure detection, whose training set contains 1327 seizures from 130 patients. Files from this dataset contain EEG-specific channels and may contain supplementary channels, such as ECG.

4.2 Data preparation

4.2.1 Data pre-processing

All signals contain acquisition noise and may have several artifacts, originated from internal or external sources (78). Artifacts originating from internal sources are due to regular physiological activity, while in external sources the causes are cable movement, electrode pop up or environmental interference.

Although in some works artifacts are not directly removed (79), their interference may be attenuated by a post processing module, which is further explained in section 4.5. In fact, the presence of a high variety of artifacts may have the potential of reducing the number of false alarms in a real product, by improving the generalization capability of an algorithm (80).

Power-line interference (50/60Hz noise), common in biosignals, may be mitigated by applying a notch filter at the power-line interference frequency (81).

The next subsections describe typical artifacts specific of ECG and EEG signals and the used strategies for attenuating their influence in epileptic signals.

4.2.1.1 ECG

ECG signals may have baseline wandering (caused by the subjects' movement or breathing) or impulsive artefacts.

In (82), a bi-directional Butterworth filter with a cut-off frequency of 3.1 Hz is applied for removing baseline wandering. Afterwards, the obtained signal is subtracted from the original.

For removing impulsive artefacts, which may impair the detection of the R peaks, a median filter is typically used. The absolute difference between the original and the median-filtered signal is computed, and divided into small segments. If the segment exceeds a certain threshold, it is considered an artifact and its values are replaced by the average of the neighbourhood (82).

4.2.1.2 EEG

Surface EEG signals have three vital and inevitable sources of artifacts, which are muscle artifacts, eyes movement/blinking and ECG signals (when it is being measured along with EEG).

In (83), an artifact rejection module is used in which large differences between the minimum and maximum values of the signal for each channel, for a given epoch, are detected. If an artifact is detected in more than 20% of the channels, the alarm is turned off for that epoch.

For denoising, a finite impulse response filter is applied in (84) for removing frequencies below 0.7Hz and above 99Hz. In (66), a band pass filter is applied for maintaining frequencies between 0.3 and 70 Hz.

4.2.2 Defining epoch length

In spectral analysis, it is considered that the frequency content of the signal is stationary over the course of time. Typically, the signals are divided into small windows (epochs), and its properties are considered to not vary significantly in that time period.

There is some variation in the definition of epoch length, in the literature. Table 4.1 shows a summary of the epoch length in state of the art works for seizure detection and prediction, for EEG, iEEG and ECG.

Table 4.1: Epoch length in state of the art works for seizure detection and prediction, for EEG, iEEG and ECG.

Signal	Detection epoch length	Prediction epoch length
EEG	2.56s (66) 3s (85)	5 (86) 30s (87)
iEEG	1s (83; 70)	4s (88) 30s (87)
ECG	30, 50 or 100 R-R intervals (89) 2s (79)	3 minutes (90) 10 minutes (91)

4.2.3 Class imbalance

As many of the datasets built for seizure detection and prediction contain long time recordings, there may be some imbalance between classes. Most classification algorithms perform optimally when classes are balanced.

For dealing with dataset imbalance, in (92), oversampling is performed for seizure detection. In (88), random subsampling is performed in an interictal vs pre-ictal classification problem. A percentage of the interictal feature samples were randomly selected for training and the remaining for testing. In (93), from EMG and accelerometer signals, windows with a standard deviation above a certain threshold were chosen as the training subset.

The effect of imbalance can also be minimized on the classifier itself. In (94), an Artificial Neural Network is used for seizure detection in EEG. However, the loss function is modified to penalize the misclassifications of the minority class (ictal) more than the majority class (interictal).

4.2.4 Channel Selection

In partial seizures, only some electrodes are required for capturing seizure activity. In many detectors present in the literature, all channels are used for feature extraction, which leads to an increase of the model complexity, computational load for training it, and to a decrease of its performance. Channel selection methods are commonly used for maintaining relevant information while mitigating these issues (79). These methodologies have been applied for EEG signals.

The simplest methodology for channel selection is using metrics such as signal variance and entropy (95). N channels with the highest metric are chosen, where N is recommended to be 4 to 6.

Alternatively, supervised approaches can be used, such as the one by (70), in which features from all channels are concatenated, and feature selection techniques combined with cross-validation are used, gradually dropping channels with a lower feature ranking.

More advanced approaches, based on deep learning, use an SSDA network for intra-channel feature extraction and channel selection (64).

4.3 Feature Extraction

The next sections describe commonly used features for seizure detection and prediction, either in classical machine problems or in threshold-based methods for classification. As there is not a perfect feature set for a given problem, the same features may be used in both seizure detection and prediction. As such, no discrimination is performed here.

4.3.1 EEG

According to the literature, features extracted from EEG can be subdivided into spectral, morphological, statistical, non linear and multivariate ((21), p. 147).

4.3.1.1 Spectral features

Features can be extracted from the frequency domain of an EEG signal. The same features may be used for EEG and iEEG recordings, as done in (86).

In (21), pp.144-147, several features are indicated for seizure detection:

(a) Signal power: power spectral density integrated in the full frequency domain;

(b) Signal bandpower: power spectral density integrated in a frequency range. For surface EEG, this may be done for each type of brainwave;

(c) Center frequency;

(d) Spectral edge frequency: the frequency below which a certain percentage of the total signal power is concentrated;

(e) Spectral entropy: flatness of the frequency spectrum;

4.3.1.2 Morphological features

Although Fourier analysis is extremely powerful for extracting the frequency content of signals, it has its limitations in seizure detection. Specifically, as epileptic spike-waveform units may retain their shape but have increasing amplitudes as the seizure progresses, this leads to a spread of the power spectrum and increases the overlap between the interictal and ictal period. As such, it is important to look at the EEG signal also in the time domain ((21), pp.144-147).

In (96), Wavelet transformation is used for seizure detection in iEEG. Wavelet transformation is a filtering process used to look at signals over different scales, by decomposing time series into components at multiple levels of resolution. Wavelets are localized in frequency and time.

The Discrete Wavelet Transform (DWT) subjects the signals to low and high pass filters, decomposing the data into approximate and detailed coefficients in various levels. The Continuous Wavelet Transform (CWT) transforms the EEG signals into scalogram plots using dilation and translation of the wavelet functions (x representing time, y representing scale, and z representing coefficient value), from which more features can be extracted (97).

Other metrics, which are amplitude related, may include peak amplitude, slope, curvature or sharpness ((21), p.147), or line length (98). Line length computes the normalized sum of amplitude changes in a sliding window, and is sensitive to changes in amplitude and frequency.

4.3.1.3 Statistical and nonlinear features

Other features of the time signal, such as statistical and nonlinear features, may also be useful.

In (21), pp.148-149, several features are mentioned:

(a) Central moments of the signal (mean, variance, skewness, kurtosis);

(b) Rank filters (minimum, median, maximum);

(c) Autocorrelation function: temporal correlation between measurements separated in time. A signal that has strong serial correlation has low frequency content;

(d) Entropy: measure of randomness;

(e) Teager energy: measure of a signal's energy that preferentially weights high-frequency activity.

In (99), the largest Lyapunov exponent, an estimation of the chaos level of the EEG data (useful for identifying transition between states) was introduced for use in both seizure prediction and detection. (100) used the Lempel-Ziv complexity metric, which measures the repeatability of episodic patterns.

4.3.1.4 Multivariate features

Multivariate features capture interactions between two or more EEG channels.

In (21), p. 149, both cross correlation and mutual information are described. Cross correlation measures the similarity of two channels as a function of the displacement of one relative to the other. Mutual information is a measure of general interdepencence between channels.

4.3.2 ECG

Typically, ECG metrics are split in time and frequency domain heart rate variability (HRV). After identifying the R peak of the ECG recordings, the RR value is calculated as the difference between consecutive R peaks, typically expressed in seconds.

4.3.2.1 Time domain heart rate variability

Heart rate variability (HRV) time-domain metrics quantify the observed variability during monitoring periods that may range from <1 min to >24h. There are optimal time periods for measuring each metric (101).

The most common HRV time-domain metrics used are (102):

(a) MeanNN: mean of RR intervals;

(b) RMSSD: root mean square of successive differences. Conventional minimum recording is 5 minutes, but periods of 10s, 30s and 60s have been proposed;

(c) SDNN: standard deviation of RR intervals. Conventional short-term recording standard is 5 minutes, but ultra-short-term recording periods from 60 to 240s are also common.

(d) NN50: number of pairs of adjacent RRI whose difference is more than 50 ms. A 2 minute epoch is recommended;

(e) PNN50: value of NN50 divided by the total number of N-N (R-R) intervals;

(f) VAR: variance of RR intervals;

4.3.2.2 Frequency domain heart rate Variability

Besides time-domain metrics, frequency domain metrics are also typically used (102):

(a) LFn: power of the low frequency band (0.04 Hz—0.15 Hz) normalized to the total power. Is typically recorded over a minimum 2 min period;

(b) HFn: power of the high frequency band (0.15 Hz—0.40 Hz) normalized to the total power. Conventionally recorded over a minimum of 1 minute;

(c) LF/HF: ratio of LF to HF, which is related to the sympathetic-parasympathetic balance of the autonomic nervous system.

For spectral analysis of heart rate variability, a sampling frequency of 4Hz was proposed for a majority of cases. This value is appropriate for studying autonomic regulation, as reliable estimates between 0 and 1Hz can be performed, which are significant frequency bands for this purpose (103).

4.3.2.3 Non-linear HRV metrics

Besides time and frequency domain features, non-linear metrics are also discriminative of seizure state (102):

(a) SD1: standard deviation of projection of the Poincare plot on the line perpendicular to the line of identity - measure of short-term variability;

(b) SD2: standard deviation of the projection of the Poincare plot on the line of identity. It is a measure of long-term variability;

(c) CSI: Cardiac Sympathetic Index, defined as the ratio between SD1 and SD2. Typically, CSI100 (over 100 R-R intervals) is calculated;

(d) CSV: Cardiac Vagal Index, defined as the logarithm of SD1 multiplied by SD2.

4.4 Model Choice and Evaluation

4.4.1 Cross validation

When dealing with time series data, traditional cross-validation should be used with care, to avoid data leakage, in which the test set contains information about segments also in the training set. This may happen since epochs in close time proximity (such as close interictal segments) may also have similar properties, and end up in different folds, one of which is used for validation (87).

For cross validation, (87) uses a percentage of later samples for validation and the remaining for training. In (82), a double-cross validation estimation is proposed, for patients with a low amount of seizures. The classifier was trained on earlier seizures, while later seizures were used for testing. In each seizure, regular k-fold cross validation is performed. In (88), leave-one-out cross validation is used with feature vectors preceding the seizure left out in each turn and tested on the remaining data.

Regular k-fold cross validation is performed in several works on seizure detection and prediction on each epoch of the signal, such as (85) and (82).

4.4.2 Dimensionality reduction

As explained in section 3.3, dimensionality reduction methods can be used for understanding the best features for patient-specific approaches and for reducing the complexity of the problem and improving computation time. Although many state of the art approaches are based on deep

learning and require no manual feature selection, this approach is followed in some cases. Table 4.2 presents a summary of the used dimensionality reduction approaches for epileptic seizure detection and prediction, for EEG, iEEG and ECG.

Table 4.2: Summary of the used dimensionality reduction approaches for epileptic seizure detection and prediction, for EEG, iEEG and ECG.

Signal	Detection dimensionality reduction	Prediction dimensionality reduction
EEG	PCA (104)	Mutual Information (66) Maximizing Class Separability (88)
iEEG	Regression Tree (105)	ReliefF (102) Maximizing Class Separability (88)
ECG	Forward Feature Selection (51)	PCA (90) ReliefF (106)

4.5 Post-processing

Due to noise, it is common that the output of a classifier fluctuates during a seizure (or during the pre-ictal period) so that it crosses the classifier's threshold. Post-processing techniques which consider how close the epochs are can be used to mitigate this.

In (87), the *k of n* method is used, which an alarm for seizure prediction is set if at least k predictions among the last n predictions were positive, using k=8 and n=10. For instance, for the last 300 seconds, an alarm is set if at least 240s led to a positive prediction. A similar methodology can be used for seizure detection.

In another work (107), a Kalman filter is used for smoothing the decision boundary.

4.6 Epileptic events classification

From selected key papers mentioned at the beginning of the chapter, the algorithm pipeline, studied population, used dataset, average performance, seizure type (in the case of ECG, where this is more relevant) and innovation are reported in the following subsections. Only the papers having different pipelines than what was previously described in Chapter 3 are further described. In appendix, tables A.1 to A.3 detail the features used in each cited paper.

Current algorithms for seizure detection achieve high sensitivity and specificity with low detection delays, but not for all seizure types. Seizure predictors have varying performance. However, the performance difference between methods cannot be directly compared in most cases, as different datasets are used with different metrics calculated.

4.6.1 EEG and iEEG

Tables 4.3 and 4.4 shows examples of significant works on seizure detection and prediction.

Table 4.3: Scalp EEG based seizure detection and prediction.

Reference	Algorithm	Population & Data	Dataset	Avg. performance	Innovation
			Seizure detection		
(85)	Pyramidal-1D-CNN	5pt. 100sz/0.6h	Bonn University	Acc: 99.1%	Useful for solutions on a chip; Model has few parameters.
(108)	Data augmentation; LSTM; Softmax classifier	5 pt. 100sz/0.6h	Bonn University	Sens: 100% Spec: 100%	Used data augmentation simulating muscle artifacts, white noise and eye blinking
(104)	EEG to spectrogram; SSDA for feature learning; SSDA-based channel selection; Softmax classifier	23 pt. 198 sz/?h	Boston	Acc: 94% ROC AUC: 0.98	Considers correlation between channels;
(109)	STFT; Integrated band-power; Adaptive thresholding	159 pt. 794 sz/25 278h	Private	Sens: 87.3% FAR: 0.22/h	Valuable for fast and effective screening of long-term scalp EEG recordings
			Seizure prediction		
(66)	Feature extraction; Feature selection (Mutual Inf.); RNN	25 pt. 86 sz/625h	Private	*Post-onset* Sens: 100% FAR: 0.023/h DL: 4 s *Pre-onset* FAR: 0.06/h DL: -51s Sens: 100% for 14/25 patients	Feasibility of seizure prediction without invasive electrodes.
(86)	Channel selection; Feature extraction; MLP classifier	227 pt. 3-33 sz per patient/162h	Private	Sens: 68 $\pm$ 22 % FAR: 0.39 $\pm$ 0.37/h	Tested different pre-ictal times.

Abbreviations: SSDA: stacked sparse denoising autoencoder; LSTM: Long short-term memory network; STFT: short-time Fourier Transform; RNN: Recurrent Neural Network; ANN: Artificial Neural Network; MLP: Multilayer Perceptron; pt: patient; sz: seizure.

Table 4.4: iEEG based seizure detection and prediction.

Reference	Algorithm	Population & Data	Dataset	Avg performance	Innovation
			Seizure detection		
(110)	Feature extraction; Random Forest classifier	12 pt 48 sz/7h	UPenn/Mayo Clinic	Sens: 91.33% DL: 3.17s Spec: 94,02%	Winning result on Kaggle competition
(70)	Channel selection; Algorithm in (110)	12 pt 48sz/7h	UPenn/Mayo Clinic	DL: 2.77s Sens: 91.95% Spec: 94.05%	Significant increase on computation efficiency
(83)	Feature extraction; SVM classifier	10 pt 67 sz/875h	Private	FAR: 0.03/h Sens: 100% (for 8/10 patients) DL: 5 s	Very good performance
(92)	Manually choose subset of electrodes; CNN applied on spectrogram	24 pt 6-92sz per pt/?h	EPILEPSIAE	Sens: 96% FAR: 10.1/h DL: 3.7s	Algorithm runs on a microchip
			Seizure prediction		
(76)	Feature extraction; Ensemble of classifiers (weighted average)	3 pt 1139sz/31 824h	Melbourne University	AUC: 85%	Winning solution for Kaggle challenge
(111)	Preprocessing: Differential window; Feature extraction (Phase Correlation); SVM Classifier; Post-processing: "k of n" alarm	18 pt 80sz/427h	Freiburg	Acc: 91.95% FAR: 2.14/pt	State of the art accuracy on this dataset
(87)	Short-time Fourier transform; CNN "k of n" alarm	*Freiburg:* 13 pt 39 sz/311.4h *Boston-MIT:* 13 pt 64sz/209h *Melbourne:* 7 pt 48sz/627.7h	Freiburg Boston-MIT (scalp) Melbourne	*Freiburg:* Sens: 81.4% FAR: 0.06/h *Boston-MIT:* Sens: 81,2% FAR: 0.16/h *Melbourne:* Sens: 75% FAR: 0.21/h	Good results in very different datasets
(88)	Feature extraction SVM classifier	18 patients 80sz/427h	Freiburg	Sens: 100% FAR: 0.0324/h	Low complexity State of the art result

Abbreviations: CNN: Convolutional Neural Network; SVM: Support Vector Machine; sz: seizure; pt: patient

Most recent EEG and iEEG seizure detection approaches cited below are based on deep learning architectures. However, machine-learning based approaches also show state of the art results

in some datasets.

Despite the differences in terms of proximity to the signal source, the average performance of seizure prediction in sEEG and iEEG has been shown to be similar (112), but this could be data-dependent.

4.6.2 ECG

Table 4.5 shows the main approaches for ECG based seizure detection. Since all datasets are private with a different population, no direct comparison can be done between the algorithms.

Despite the promising results that are found for ECG-based seizure prediction, most of the proposed prediction algorithms have reproducibility problems in other datasets, and a lack of prospective validation studies (113). The overall quality of studies on seizure detection using autonomic parameters (such as ECG) is low, due to small population sizes and high study heterogeneity. Current studies have a lack of long-term and real-time ambulatory monitoring.

Most of the presented approaches use standard HRV-related feature extraction followed by a classifier, and may have a dimensionality reduction step in between. However, some works, which will be focused here, have relevant differences.

As the ictal HR patterns may take over 2 minutes in some cases, leading to a too large detection delay, (114) analyses only the linear phase of ictal HR patterns, denoted as HRI. Compared to previous approaches, the detection delay and false positive rate are reduced.

In (90), seizure prediction is modeled using Multivariate Statistical Process Control, a statistical approach which models the correlation among variables with PCA. PCA finds linear combinations of variables which describe trends in a dataset, and points that do not follow this trend are considered to be anomalies.

Table 4.5: ECG-based seizure detection and prediction studies.

Reference	Algorithm	Population	Types of sz.	Avg performance	Innovation
			Seizure detection		
(51)	HRI detection Feature extraction (from HRI period); Feature selection (forward selection); Patient-independent SVM classifer	17 pt (TLE) 127sz/918h	FOIA, FOBTC	Sens: 81.9 % FAR: 1.97/h DL: 17.8 s	Less false alarms compared to state of art High DL
(114)	Patient-independent classifier (51); Transfer learning to create patient dependent classifier	6 pt (TLE) 74sz/206h	FOIA	Sens: 89.8% FAR: 1.1/h DL: 11.7 s	Personalized alarms
(115)	Patient-specific adaptive classifier (51)	19 pt (TLE) 153sz/2883h	GOS, FOS	Sens: 77.6% FAR: 2.56/night	Adapts to signal changes
(116)	CSI100 value thresholded	5 pt (TLE) 11sz/?h	FOIA	Sens: 76% DL: 16 s	Differentiates ictal and exercise-induced changes
(117)	Feature extraction; Clustering for classification	37 pt 98sz/?h	FOS,GOS	Sens: 100% (FOS) Sens: 90% (GOS)	Uses ECG morphology (PCA components)
			Seizure prediction		
(90)	Feature extraction; Multivariate Statistical Process Control (anomaly detection)	8 pt 22sz/18.9h	Not specified	Sens: 91% FAR: 0.7/h SPH: 10s	Models seizure prediction as an anomaly detection problem
(102)	Feature extraction; Feature selection (ReliefF); SVM Classifier	15 pt 38sz/?h	Not specified	Sens: 89.06%; FAR: 0.41/h SPH: 15 min	Showed the feasibility of a long prediction window using ECG
(118)	Feature extraction; Adaptive thresholding on HRV features	16 pt 65sz/660h	Not specified	Sens: 78.59% SPH: 110s FAR: 0.21/h	Simple algorithm which shows potential for seizure prediction.
(91)	Feature extraction; Feature selection (PCA); SVM Classifier	With epilepsy: 12 pt (TLE) 34sz/55.2h Healthy: 6 pt 0sz/123.6h	FOIA	Epileptic: Sens: 94,1% FAR: 0.49/h Healthy: FAR: 0.19/h	Tested the algorithm on long-term ECG for healthy people.

Abbreviations: FOIA: focal onset impaired awareness; FOBTC: focal onset to bilateral tonic–clonic; GOS: generalized onset seizure; FOS: focal onset seizure; TLE: Temporal Lobe Epilepsy; SPH: Seizure Prediction Horizon; sz: seizure; pt: patient.

4.6.3 Multimodal approaches

Unimodal detectors still have an unacceptable amount of false alarms (119). Multimodal algorithms might help lower this false alarm rate, as different seizure types may lead to alterations in only some biosignals - for instance, a non-motor seizure shows no motor changes, and as such, a seizure detector based on EMG or accelerometer would not work for these types of seizures. In the literature, a 30-57% reduction of the false alarm rate was observed, using a multimodal approach (93).

Table 4.6 shows a few examples of multimodal seizure detection algorithms. Missing are many EDA and accelerometer based algorithms, which are not described here due to such sensors not being used for the dissertation that will follow this monograph.

The major difference between multimodal methods is the way the biosignals are integrated with each other for achieving the final result, such as late integration (79; 120) or early integration (121).

The closest approach to what is proposed in this dissertation is (122). On an EMU, a full channel monitoring using EEG, HR, SpO2 and EDA is performed. The performance of seizure detection using HR, SpO2, EDA is compared to three-channel EEG. If it is high enough using the non-EEG approach, the patient is sent home with a wrist-device. Alternatively, a hybrid device containing all the signals is used. The algorithm is based on detecting trends in the mean value of parameters, in a given window, according to known physiological changes.

Table 4.6: Examples of multimodal seizure detection algorithms.

Reference	Sensor	Algorithm	Population & Data	Avg performance	Innovation
		Seizure detection			
(120)	EEG EMG ECG	*EEG:* Rhythmic pattern detection compared to average spectrum; *EMG:* Line Length compared to baseline; *ECG:* CSI100 over 100 beats; Threshold; Late fusion (*OR*)	92 pt 494sz/11 978h	Sens: 86% FAR: 0.7/h	High sensitivity using reduced electrode montages.
(121)	EDA Accelerometer	*Acc:* time/frequency domain features non-linear features; *EDA:* time domain features; SVM classifier (early fusion).	80 pt (7 with sz.) 16sz/4,213h	FAR: 0.03/h Sens: 94%	Novel methodology for generalized tonic–clonic sz. detection.
(122)	*First stage:* HR, SpO2 and EDA *Second stage:* HR, SpO2, EDA, ECG *Third stage:* EEG	*First stage:* Search for increases in HR and EDA and decreases in SpO2 *Second stage:* Personalized parameters for each patient *Third stage (for some subjects):* EEG-feature extraction; 3-Channel selection based on best features; kNN classifier.	10 pt 26sz/339h	*First stage:* Sens: 100% (7/10 pat., 10 sz) *Second stage:* Sens: 100% (6/10 pat., 11sz) *Third stage:* Sens: 100% (2/3 pat)	Wearable approach concerning a patient-specific biosignal selection
(93)	Accelerometer EMG ECG	*ECG:* time domain features; *Acc.:* frequency and time domain features; *EMG:* time domain features; SVM classifier; Logical OR on both alarms, in a certain time window	7 pt 22 sz/224h	ECG+Acc. (best) Sens: 90,9% FAR: 0.08/h	Multimodal sensors lead to 75% less false alarms; ACM changes are slower than EMG;

Abbreviations: SpO2: Oxygen Saturation; sz: seizure; Acc: Accelerometer; SVM: Support Vector Machine; CSI100: Cardiac Sympathetic Index;

4.6.4 Existing solutions in the market

Besides academic solutions, a research for solutions for epileptic seizure detection and prediction already available in the market are presented in Table 4.7. These were found by combining a patent search performed by INESC TEC's Patent Office for epileptic seizure detection and prediction patents, and a review article found in the literature (123). Solutions which have already been removed from the market were excluded.

Table 4.7: Existing solutions for epileptic seizure detection and prediction in the market.

Sensor(s)	Brand	Main Features	Seizure type(s)
Seizure detection			
HR ACM	Epiwatch	Data collected by iPhone / Apple Watch	Tonic-clonic Generalized tonic Hyperkinetic Clusters of short myoclonic/tonic seizures
PPG ACM	NightWatch	Arm band	Night time motor seizures
ACM	Emfit Seizure Monitor	Bed motion sensor with accelerometer under mattress	Night time motor seizures
ACM	EpiCare	Armband with acceleration sensors	Tonic-clonic
ACM	MP5, MP2, ST2	Bed motion sensor and vocalization microphone	Tonic-clonic
ACM EDA TEMP	Embrace Affectiva Q-Sensor	Wristband for seizure detection	Generalized tonic-clonic
EMG	Brain Sentinel	Device worn with strap on biceps	Generalized tonic-clonic
EEG	Epihunter	Strap with two electrodes placed on frontal region	Absence seizures
Seizure prediction			
EEG	Epistemic	Two electrodes placed on the forehead	Not detailed

There are not many devices for seizure prediction in the market, as they are still in initial studies. Seizure detection devices are mostly focused on motor seizures, leading to opportunities for further development in this area.

4.7 Conclusion

After analysing the state of the art, it is possible to infer about the research gaps in specific areas. Note that all comments here are performed to the best of our knowledge, as some relevant papers could have been missed during the state of the art research.

Although deep learning methodologies are commonly used in ECG for biometrics, arrhythmia detection and heartbeat classification, among others (59) there are no studies regarding its performance in epileptic seizure detection and prediction. There is also a lack of deep learning-based studies for multimodal approaches.

There are also no studies based on multimodal seizure prediction (e.g. using iEEG/EEG and ECG data).

Focusing on data availability, there are no public datasets for ECG signals during the pre-ictal and ictal period, which makes it difficult for other researchers to test their algorithms on openly available data.

Chapter 5

Epileptic Seizure Detection

This chapter shows a proof of concept on 8 patients of a personalized algorithm which uses electrocardiogram (ECG) and EEG for epileptic seizure detection.

5.1 Dataset

In a joint research line between University of Porto and the Ludwig Maximillian University of Munich, we have been working on 3D Video-EEG, building the largest multimedia database of this kind. Currently, the database holds more than 300 seizures from over 100 patients (124). Besides RGB video and 3D infrared radar, we store long term surface EEG and ECG signals synchronously recorded at 256Hz and stored. Three minutes post seizures were removed from this dataset as this is not representative of natural monitoring conditions, as it is often the period where the patient is recovering from the seizure.

The rest of the dataset, also at 256Hz, was taken from Temple University Hospital (77) (TUH-EEG Corpus), an ongoing data collection effort that has recently released a training set containing 1327 seizures from 130 patients, of which the irrelevant signal parts were manually removed by healthcare professionals. TUH data was chosen since the data available from Munich at the time of the dissertation was not enough to validate an algorithm. Also, this way, the developed algorithms can be tested on publicly available data.

A subset of 8 patients from both Munich and TUH-EEG was chosen to represent different seizure types, lengths and recording durations, further detailed in Table 5.1. A total of around 93 seizures and 110 hours of data was analyzed.

5.2 Data processing

The following subsections describe the transformations applied to the data as a way of preprocessing it, to reduce the irrelevant information and computation time.

Table 5.1: Detailed information of the 8 patients used for this study, including the number of seizures, its duration, the imbalance ratio between classes, epilepsy syndrome and seizure type(s).

Patient ID	**Sz. amount**	**Duration**	**Interictal/ictal ratio**	**Syndrome**	**Seizure type(s)**
MUN1	19	37h	182	Temporal Lobe Epilepsy	Hypermotor to generalized tonic-clonic Bilateral tonic to automotor
MUN2	11	24h04m	759	Paracentral Epilepsy	Aura to Tonic Clonic Seizure
MUN3	6	47h40m	1037	Frontal Lobe Epilepsy	Aura to Complex Motor
TUH1	8	21m41s	1.6	Unknown	Tonic-clonic seizures
TUH2	11	22m43s	19.7	Frontal Lobe Epilepsy	Absence seizures
TUH3	17	24m54s	2.17	Occipital Lobe Epilepsy	Complex partial seizures
TUH4	12	12m25s	0.58	Temporal Lobe Epilepsy	Generalized seizures
TUH5	9	58m43s	18.0	Frontal Lobe Epilepsy	Status Epilepticus
Total	93	110h48m46s			

Initially, EEG and ECG data was split in three seconds segments. Each segment was labeled as ictal (seizure period) or interictal (non-seizure period) depending on the majority of samples of each class.

No data filtering or artifact rejection was applied, as it has been stated in the literature that artifacts may improve the generalization capability of an algorithm (80). Moreover, seizures itself may cause artifacts due to cable movement of sudden electrode removal.

5.2.1 Channel Selection

Channel selection can significantly reduce the computational cost of the algorithm while also decreasing the amount of irrelevant information the classifier needs to learn, as not all channels contain epileptic activity. In EEG recordings, if no information regarding seizure location was available, the four channels with highest variance during the ictal period were selected, as previously recommended in the literature (95). However, if healthcare professionals had previously done the task of labeling the channels involved in seizures (or the epilepsy syndrome), manually selected channels from that region were used, to decrease the source of errors.

This algorithm of channel selection was validated by comparing the automatically selected channels were compared with the affected areas (Table 5.2).

Table 5.2: Validation of the channel selection algorithm

Patient ID	**Automatic channel selection**	**Affected areas/channels**
MUN1	FP1, FP2, P3, F12	FT10, T8, TP10, FT8
MUN2	FT8, F4, F12, P12	C4, P4, CZ, C3
MUN3	FP1, FT10, AF8, TP10	Fronto-central/Frontal
TUH1	C3, C4, T4, CZ	Unknown
TUH2	FP2, F3, F4, FZ	Frontal
TUH3	FP2, F3, C3, P3	Occipital
TUH4	F3, A2, P4, PZ	Temporal
TUH5	FP1, F3, F7, T1	Frontal

From the Munich dataset, there was information regarding the epileptic activity in each electrode. However, the TUH cases only have information regarding the seizure onset zone, and do not consider the possibility of the spread of epileptic activity to other channels during a seizure.

Channels containing the highest variance may have some artifacts non-seizure derived, such as eye blinking. This could have happened in some cases, since FP1 and FP2, channels subjected to eye blinking artifacts, are typically selected by this algorithm.

For MUN3, TUH2 and TUH5, the automatically selected channels seem representative of the affected areas. For TUH1 the affected areas were not disclosed, but a lot of channels from the same area are being selected. As such, epileptic activity could be more discriminating on the central brain area.

When possible, manual channel selection was used, selecting all channels in affected areas and the ones closest to it. This was possible for all patients except TUH1, which decreases the error regarding channel selection, by maintaining the channels that were clinically validated. Automatic channel selection seems viable for some patients, and was used in cases where no information regarding the electrodes containing epileptic activity is seen (TUH1).

5.3 Classification

For classification of seizure state, three classifiers were tested: based on Convolutional Neural Networks, on Fully Convolutional Networks and on Long Short-Term Memory networks. The networks' architectures are further described in the following subsections.

5.3.1 Convolutional Neural Network

A simple Convolutional Neural Network was built, with an architecture that allowed for fast training and did not have an excessively high number of parameters. The CNN has two convolutional layers, with kernel size 7, containing 6 and 12 filters, respectively. Each convolutional layer is followed by a sigmoid activation function, a Dropout layer with a rate of 50%, for regularization, and an average pooling layer. A sigmoid classifier is then fully connected to the last layer's output. The total number of trainable parameters is 1549, in ECG (single channel), and 1675 in EEG (four channels). Figure 5.1 represents the network's architecture.

5.3.2 Fully Convolutional Networks

Fully Convolutional Neural Network (FCNs) is a network that was recently proposed for classifying time series (125), but was not directly applied for seizure detection. It was chosen here due to their efficacy for time series classification in several benchmark datasets.

The used network contains three convolutional blocks, each with three operations: convolution, batch normalization (BN) and ReLu activation. Then, a global average pooling (GAP) is applied, averaging the results of the last convolutional block. A sigmoid classifier is then fully connected to GAP's output.

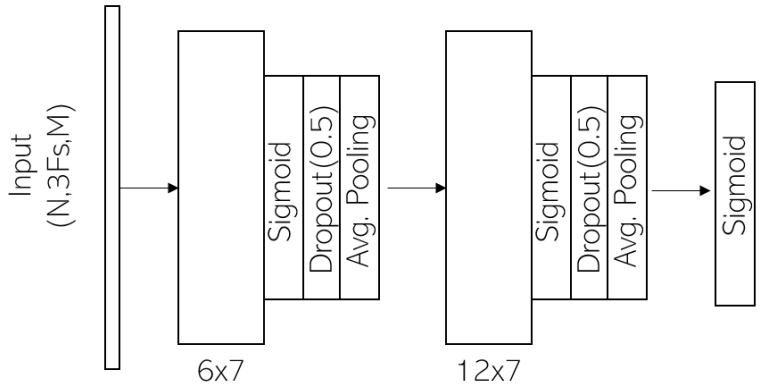


Figure 5.1: Architecture of the Convolutional Neural Network. In the input layer, N is the number of epochs used for training, 3Fs is the sampling frequency multiplied by three seconds, and M is the number of channels (1 for ECG and 4 for EEG).

Regarding filter size, the convolution's stride is equal to 1, with zero padding applied. The convolutions contain 128, 256 and 128 filters, respectively, with filter lengths equal to 8, 5 and 3.

The total number of trainable parameters is 264 833, in ECG (single channel), and 267 905 in EEG (four channels). Figure 5.2 shows a representation of the network used.

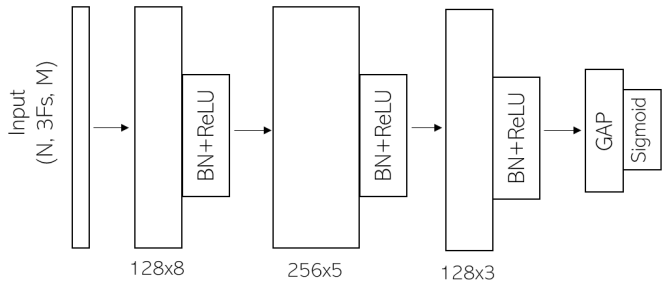


Figure 5.2: Architecture of the Fully Convolutional Neural Network, extracted from (125). In the input layer, N is the number of epochs used for training, 3Fs is the sampling frequency multiplied by three seconds, and M is the number of channels (1 for ECG and 4 for EEG).

5.3.3 Long Short-Term Memory

The used LSTM architecture was previously proposed for seizure detection in Bonn dataset (108). The data samples are input into a fully connected LSTM layer of 100 neurons, as a way of learning the short and long term dependencies between segments in each signal, and returning hidden state output for each input time step (with the same size as the input). The interest of such an architecture in this context is harnessing the LSTMs' ability of learning and extracting features across the time steps in a window of 3 seconds, and not across windows (stateless LSTM).

Afterwards, a time-distributed Dense layer of 50 units is applied, to calculate the cost function on all timesteps.

Finally, the output of the dense layer is passed through a 1D global average pooling layer, which is then input into a final softmax/sigmoid layer. The total number of trainable parameters is 45 901, in ECG (single channel), and 47 101 in EEG (four channels). The network is schematically represented in Figure 5.3.

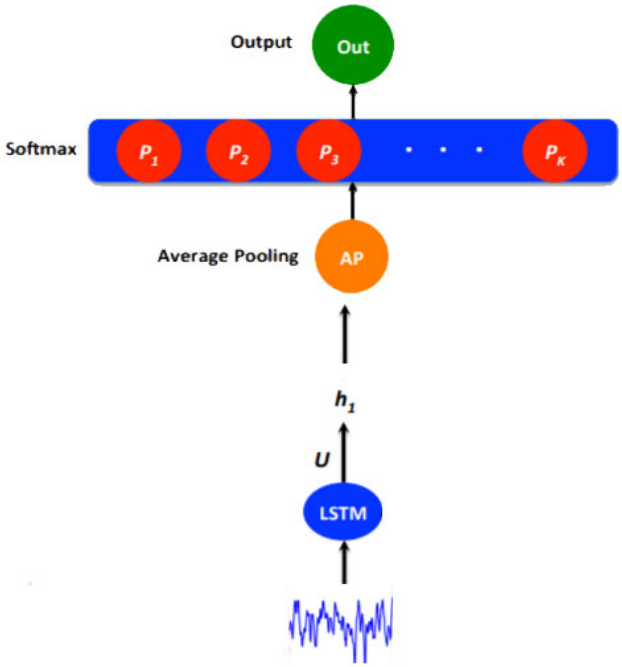


Figure 5.3: Architecture of the used LSTM Network, with a single epoch input. Figure adapted from (108). U is the output of the LSTM layer, with the same size as the input, and h1 is the dense layer.

5.3.4 Cross-validation strategy

Initially, to identify groups of epochs containing seizures, the region labeling algorithm (126) was used in the ground truth. This algorithm, visually explained in Figure 5.4, iterates through the whole timeseries (in this case, composed of zeros and ones), starting with zero regions. When a new non-zero element is found, the region index is incremented. If that element has a neighbor that is not the background and has already been labeled with a certain region, the same region as the neighbor is given. As such, epochs belonging to the same seizure are grouped into the same region.

Epochs were grouped from the end of a seizure to the end of the seizure following it, according to the labeled regions (Fig. 5.5). As such, each epoch contained the period before a seizure and the seizure period. During training, a grouped leave one out cross validation was performed.

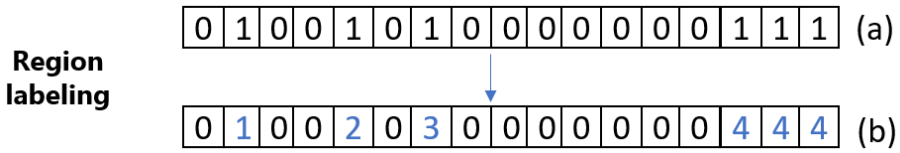


Figure 5.4: Region labeling algorithm applied to signal epochs.

This strategy guarantees that temporally close epochs from the same seizure are not in the train and test folds, which would give an inflated estimate of model performance (allowing for a possible data leakage). In each fold of the training set, the classifier's weights were set to random and the interictal period was downsampled to be eight times the ictal period, decreasing the effect of class imbalance. This value was determined iteratively, by cross validation on a few seizures from all Munich patients. In the cases where the interictal to ictal ratio was not higher than 8 (see Table 5.1), downsampling was performed so that the ratio between both classes was equal. In the validation set, the original size segment was used. This approach realistically simulates a production environment, as the model is intended to be trained in the Epilepsy Monitoring Unit on available seizure data, and should have similar performance for detecting other seizures in a longer period, where there's a higher class imbalance.

Data augmentation was applied to each training epoch as a way to increase training data: random noise with a standard deviation of 0.05 and mean 0 was added, and a random scaling with a normal distribution having standard deviation of 0.01 and mean 0 was applied. These values were selected as they did not lead to significant changes in the signal's shape (by visual analysis) and led to a better model convergence during training, during an initial analysis. More data augmentation strategies will be studied for the best model, but were not considered in this initial phase due to the significant increase in computation time.

The network was trained for a maximum of 100 epochs or after the loss stopped decreasing by a factor of 0.05 for 40 epochs. Adam optimizer (127) was used with an initial learning rate of 0.001 for FCN and 0.01 for LSTM/CNN, determined iteratively after a few cross validation folds, on a reduced number of patients. The model with the highest geometric mean between specificity and sensitivity in the validation set was then chosen as the best model. Due to class imbalance, classes were weighted accordingly to the ratio of interictal to ictal samples, and the binary cross-entropy loss was used.

The output of the classifier was post-processed as shown in Fig 5.6. In this algorithm, the initial predicted sequence is shown in (a). Initially, an algorithm which changes all elements with more than one non-null neighbor to non-null is applied (b). Afterwards, the output (c) is considered to be ictal (class 1) if N consecutive epochs were classified as ictal, where N is patient-dependent, as different seizure types have different lengths and patterns. N was picked in order to maximize the number of detected seizures. Depending on how much of the signal we need to see before making the optimal decision, the detection latency in a real time algorithm varies. As N is picked according to the cross-validated output, this decreases the additional bias of the classifier.

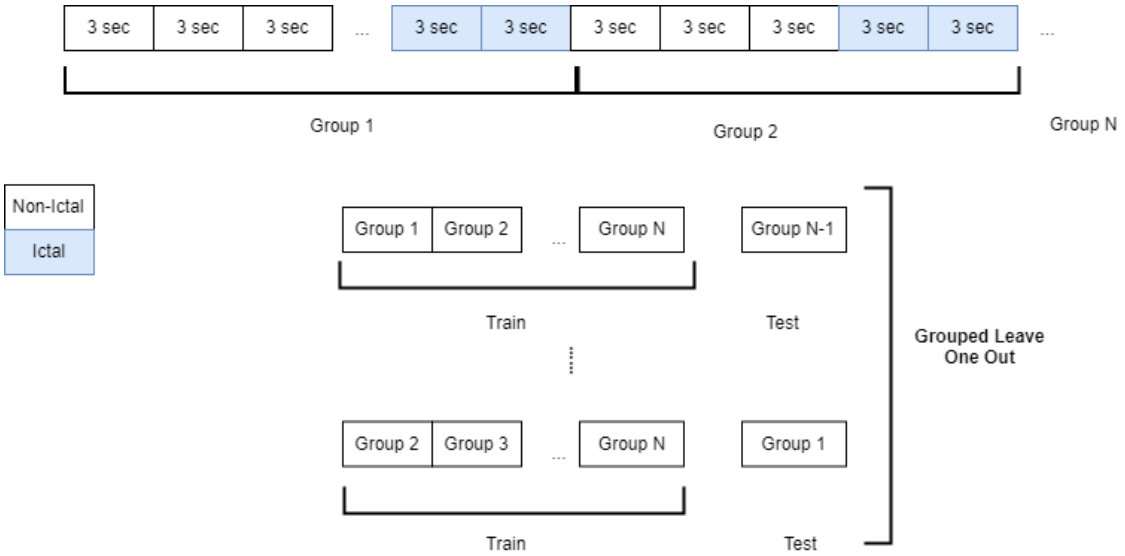


Figure 5.5: Cross validation strategy for seizure detection and prediction.

The true amount of false positives is then calculated according to the amount of regions (after region labeling is applied to the post processed output) that were wrongly predicted, as shown in the figure. This way, the false positive rate is not calculated by each epoch, but by group, as false alarms that are subsequent are not considered in real life conditions.

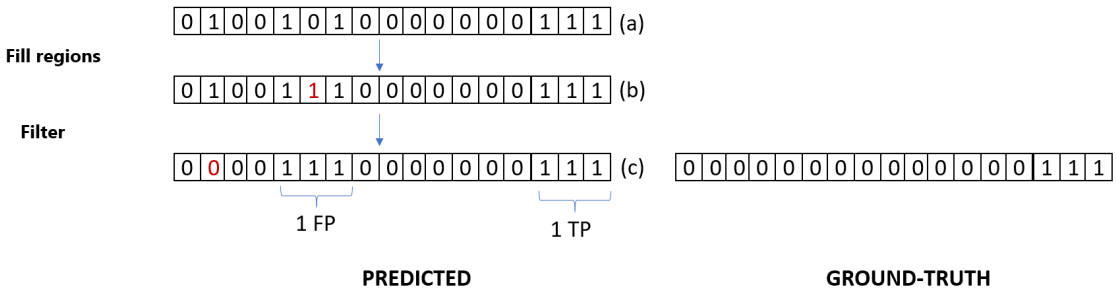


Figure 5.6: Post processing strategy for seizure detection.

To obtain a multimodal approach, the logical AND operator was used in the output of all classifiers after post-processing (late integration approach).

5.4 Results

The following subsections show the obtained classification results.

5.4.1 Comparison between classifiers

Table 5.3, 5.4 and 5.5 summarize the obtained results for each patient, for the FCN, LSTM and CNN architectures, respectively. As the seizure group duration is very variable and often low (less

Table 5.3: Summary of the results for EEG, ECG and multimodal (MM) fusion (FCN).

Patient	EEG	ECG	MM	Alarm size (epochs)
MUN1	**Sens/sz:** 1.00 **Sens/epoch:** 0.89 ± 0.12 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 1.6 ± 3.4 **FPR/h:** 2.6 ± 2.2 **Norm. FPR:** 5.06 ± 4.3	**Sens/sz:** 0.79 **Sens/epoch:** 0.56 ± 0.35 **Spec/epoch:** 0.98± 0.02 **DL (s):** 3.0 ± 5.2 **FPR/h:** 4.2 ± 3.5 **Norm. FPR:** 8.18 ± 6.8	**Sens/sz:** 0.79 **Sens/epoch:** 0.54 ± 0.34 **Spec/epoch:** 1.00 ± 0.0 **DL (s):** 4.8 ± 8.9 **FPR/h:** 0.5 ± 0.7 **Norm. FPR:** 0.97 ± 1.4	2
MUN2	**Sens/sz:** 0 **Sens/epoch:** 0 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** - **FPR/h:** 0.86 ± 1.3 **Norm. FPR:** 1.88 ± 2.8	**Sens/sz:** 0.63 **Sens/epoch:**0.55 ± 0.44 **Spec/epoch:** 0.94 ± 0.10 **DL (s):** 1.7 ± 2.2 **FPR/h:** 16.2 ± 30.8 **Norm. FPR:** 35.3 ± 67.2	**Sens/sz:** 0 **Sens/epoch:** 0 **Spec/epoch:** 1.0 ± 0 **DL (s):** - **FPR/h:** 0.2 ± 0.6 **Norm. FPR:** 0.4 ± 1.3	1
MUN3	**Sens/sz:** 0.33 **Sens/epoch:** 0.23 ± 0.36 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 4.5 ± 1.5 **FPR/h:** 1.3 ± 2.8 **Norm. FPR:** 10.3 ± 22.2	**Sens/sz:** 0.67 **Sens/epoch:** 0.63± 0.45 **Spec/epoch:** 0.93 ± 0.05 **DL (s):** 0.0 ± 0.0 **FPR/h:** 16.6 ± 10.5 **Norm. FPR:** 131.9 ± 83.4	**Sens/sz:** 0.17 **Sens/epoch:** 0.1 ± 0.22 **Spec/epoch:** 1.0 ± 0 **DL (s):** 9 ± 0 **FPR/h:** 0.04 ± 0.1 **Norm. FPR:** 0.32 ± 0.79	1
TUH1	**Sens/sz:** 1.00 **Sens/epoch:** 1.0 ± 0.00 **Spec/epoch:** 0.95 ± 0.06 **DL (s):** 0.0 ± 0.0	**Sens/sz:** 1.00 **Sens/epoch:** 0.98 ± 0.03 **Spec/epoch:** 0.97 ± 0.05 **DL (s):** 1.25 ± 2.2	**Sens/sz:** 1.00 **Sens/epoch:** 0.98 ± 0.03 **Spec/epoch:** 1.0 ± 0.00 **DL (s):** 1.25 ± 2.2	2
TUH2	**Sens/sz:** 0.9 **Sens/epoch:** 0.9 ± 0.3 **Spec/epoch:** 0.74 ± 0.29 **DL (s):** 0	**Sens/sz:** 0.9 **Sens/epoch:** 0.8 ± 0.3 **Spec/epoch:** 0.83 ± 0.28 **DL (s):** 0 ± 0	**Sens/sz:** 0.91 **Sens/epoch:** 0.9 ± 0.3 **Spec/epoch:** 0.86 ± 0.29 **DL (s):** 0 ± 0	1
TUH3	**Sens/sz:** 1.00 **Sens/epoch:** 0.90 ± 0.14 **Spec/epoch:** 0.96 ± 0.12 **DL (s):** 2.06 ± 3.8	**Sens/sz:** 0.88 **Sens/epoch:** 0.63 ± 0.31 **Spec/epoch:** 0.85 ± 0.27 **DL (s):** 2.9 ± 3.03	**Sens/sz:** 0.88 **Sens/epoch:** 0.62 ± 0.30 **Spec/epoch:** 1.00 ± 0.00 **DL (s):** 4.2 ± 4.09	2
TUH4	**Sens/sz:** 1.00 **Sens/epoch:** 0.83 ± 0.21 **Spec/epoch:** 0.92 ± 0.18 **DL (s):** 1.2 ± 1.5	**Sens/sz:** 1.00 **Sens/epoch:** 0.79 ± 0.26 **Spec/epoch:** 0.92 ± 0.16 **DL (s):** 0.9 ± 1.4	**Sens/sz:** 1.00 **Sens/epoch:** 0.68 ± 0.3 **Spec/epoch:** 1 ± 0 **DL (s):** 1.2 ± 1.5	2
TUH5	**Sens/sz:** 1.00 **Sens/epoch:** 0.98 ± 0.06 **Spec/epoch:** 0.95 ± 0.07 **DL (s):** 0.0 ± 0.0	**Sens/sz:** 1.00 **Sens/epoch:** 1.0 ± 0.0 **Spec/epoch:** 0.91 ± 0.19 **DL (s):** 0 ± 0	**Sens/sz:** 1.00 **Sens/epoch:** 0.98 ± 0.06 **Spec/epoch:** 0.98 ± 0.03 **DL (s):** 0.0 ± 0.0	2

than 10 minutes of recordings) in TUH patients, false positive rate was not calculated. Specificity is a relative value and can give a better estimate of the performance. The sensitivity per epoch and per seizure, the detection latency and the number of consecutive epochs required to trigger a seizure alarm are also reported (hereupon referred as alarm size).

Moreover, a new metric not found in the state of the art is proposed (Norm. FPR). In seizure detectors, the importance of false alarm rate is dependent on the number of seizures the patient has per day. Norm. FPR is proposed as the false alarm rate normalized by the average seizure frequency (seizures per hour).

For instance, if the false positive rate is high and seizure frequency is low, this metric is lower, meaning that false positives are not as important, since the patient does not need a close control. On the other hand, with a low false positive rate and high seizure frequency, this metric is higher. If they are both on the same scale, the metric is closer to 1.

Table 5.4: Summary of the results for EEG, ECG and multimodal (MM) fusion (LSTM).

Patient	EEG	ECG	MM	Alarm size (epochs)
MUN1	**Sens/sz:** 1.00 **Sens/epoch:** 0.96 ± 0.08 **Spec/epoch:** 0.98 ± 0.02 **DL (s):** 1.1 ± 2.9 **FPR/h:** 4.8 ± 3.5 **Norm. FPR:** 9.3 ± 6.8	**Sens/sz:** 1.00 **Sens/epoch:** 0.76 ± 0.21 **Spec/epoch:** 0.92 ± 0.08 **DL (s):** 3.6 ± 5.2 **FPR/h:** 15.8 ± 16.2 **Norm. FPR:** 30.8 ± 31.5	**Sens/sz:** 1.00 **Sens/epoch:**0.72 ± 0.21 **Spec/epoch:** 0.99 ± 0.01 **DL (s):**4.1 ± 5.3 **FPR/h:** 2.1 ± 1.7 **Norm. FPR:** 4.1 ± 3.3	2
MUN2	**Sens/sz:** 1.00 **Sens/epoch:** 0.98 ± 0.07 **Spec/epoch:** 0.91 ± 0.04 **DL (s):** 0.0 ± 0.0 **FPR/h:** 30.0 ± 12.1 **Norm. FPR:** 65.5 ± 26.4	**Sens/sz:** 1.00 **Sens/epoch:**0.91 ± 0.11 **Spec/epoch:** 0.83 ± 0.15 **DL (s):** 1.1 ± 1.4 **FPR/h:** 39.9 ± 34.6 **Norm. FPR:** 87.1 ± 75.5	**Sens/sz:** 0.91 **Sens/epoch:** 0.82 ± 0.29 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 1.2 ± 1.5 **FPR/h:** 4.3 ± 3.8 **Norm. FPR:** 9.4 ± 8.3	1
MUN3	**Sens/sz:** 1.00 **Sens/epoch:** 0.67 ± 0.29 **Spec/epoch:** 0.97 ± 0.02 **DL (s):** 7 ± 11.8 **FPR/h:** 7.5 ± 4.6 **Norm. FPR:** 59.6 ± 36.5	**Sens/sz:** 0.83 **Sens/epoch:** 0.69 ± 0.35 **Spec/epoch:** 0.97 ± 0.02 **DL (s):** 2.4 ± 2.25 **FPR/h:** 8.5 ± 6.2 **Norm. FPR:** 67.5 ± 49.3	**Sens/sz:** 0.83 **Sens/epoch:** 0.53 ± 0.31 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 3.0 ± 4.6 **FPR/h:** 1.3 ± 1.5 **Norm. FPR:** 10.3 ± 11.9	3
TUH1	**Sens/sz:** 1.00 **Sens/epoch:** 0.97 ± 0.06 **Spec/epoch:** 0.97 ± 0.07 **DL (s):** 1.25 ± 2.7	**Sens/sz:** 1.00 **Sens/epoch:** 0.95 ± 0.08 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 2.5 ± 3.6	**Sens/sz:** 1.00 **Sens/epoch:** 0.93 ± 0.10 **Spec/epoch:** 1.0 ± 0.00 **DL (s):** 2.75 ± 3.9	3
TUH2	**Sens/sz:** 0.91 **Sens/epoch:** 0.91 ± 0.29 **Spec/epoch:** 0.78 ± 0.27 **DL (s):** 0.0 ± 0.0	**Sens/sz:** 0.91 **Sens/epoch:** 0.91 ± 0.29 **Spec/epoch:** 0.86 ± 0.15 **DL (s):** 0.0 ± 0.0	**Sens/sz:** 0.91 **Sens/epoch:** 0.91 ± 0.29 **Spec/epoch:** 0.86 ± 0.28 **DL (s):** 0.3 ± 0.9	1
TUH3	**Sens/sz:** 1.00 **Sens/epoch:** 0.77 ± 0.22 **Spec/epoch:** 0.88 ± 0.15 **DL (s):** 5.7 ± 12	**Sens/sz:** 0.76 **Sens/epoch:** 0.47 ± 0.32 **Spec/epoch:** 0.85 ± 0.22 **DL (s):** 6.4 ± 4.9	**Sens/sz:** 0.65 **Sens/epoch:** 0.36 ± 0.3 **Spec/epoch:** 0.98 ± 0.04 **DL (s):** 6.5 ± 4.09	2
TUH4	**Sens/sz:** 0.9 **Sens/epoch:** 0.73 ± 0.39 **Spec/epoch:** 0.95 ± 0.15 **DL (s):** 1.0 ± 2	**Sens/sz:** 0.8 **Sens/epoch:** 0.64 ± 0.38 **Spec/epoch:** 0.81 ± 0.32 **DL (s):** 1.13 ± 2.1	**Sens/sz:** 0.7 **Sens/epoch:** 0.37 ± 0.37 **Spec/epoch:** 1.00 ± 0.00 **DL (s):** 2.1 ± 2.6	2
TUH5	**Sens/sz:** 1.00 **Sens/epoch:** 0.95 ± 0.13 **Spec/epoch:** 0.97 ± 0.05 **DL (s):** 0 ± 0	**Sens/sz:** 0.88 **Sens/epoch:** 0.86 ± 0.31 **Spec/epoch:** 0.92 ± 0.1 **DL (s):** 0 ± 0	**Sens/sz:** 0.88 **Sens/epoch:** 0.84 ± 0.3 **Spec/epoch:** 0.99 ± 0.022 **DL (s):** 0 ± 0	2

On average, high sensitivity values, similar to the state of the art, are presented here for detecting seizures. For classification per epoch, sensitivity is more variable. However, what is most important is detecting a seizure and triggering an alarm, which may prevent life-threatening conditions, and not estimating its full duration. Specificity (and false positive rate) shows an increase in multimodal approaches, on average, but with a decrease in sensitivity, and a higher variability in sensitivity per epoch, in some cases. It is preferable to have high sensitivity in difficult cases, with motor symptoms which may lead to injuries, such as tonic clonic seizures.

The architecture based on LSTM shows more potential for detecting seizures, as all patients have sensitivities per seizure above 76%, while for FCN the minimum value is 33% and for CNN 45%. As such, LSTM was chosen to be the best classifier for this subset, and its results are further described.

For patient TUH1, both EEG and ECG approaches showed relevant results as unimodal detectors. Multimodal detection did not lead to significant improvements. These results are to be expected since the patient has violent tonic-clonic seizures, easily detected using either EEG or ECG. This patient could benefit from a single wearable, such as a non intrusive ECG device,

Table 5.5: Summary of the results for EEG, ECG and multimodal (MM) fusion (CNN).

Patient	EEG	ECG	MM	Alarm size (epochs)
MUN1	**Sens/sz:** 1.00 **Sens/epoch:** 0.95 ± 0.09 **Spec/epoch:** 0.96 ± 0.03 **DL (s):** 1.1 ± 2.9 **FPR/h:** 8.9 ± 7.3 **Norm. FPR:** 17.3 ± 14.2	**Sens/sz:** 0.95 **Sens/epoch:** 0.41 ± 0.14 **Spec/epoch:** 0.93 ± 0.15 **DL (s):** 5.3 ± 8.6 **FPR/h:** 13.9 ± 18.4 **Norm. FPR:** 27.1 ± 35.8	**Sens/sz:** 0.95 **Sens/epoch:**0.41 ± 0.15 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 6.1 ± 9.6 **FPR/h:** 3.6 ± 3.9 **Norm. FPR:** 7.0 ± 7.6	2
MUN2	**Sens/sz:** 0.91 **Sens/epoch:** 0.82 ± 0.31 **Spec/epoch:** 0.80 ± 0.13 **DL (s):** 0.9 ± 1.9 **FPR/h:** 58.6 ± 31.3 **Norm. FPR:** 127.8 ± 68.2	**Sens/sz:** 0.45 **Sens/epoch:**0.33 ± 0.40 **Spec/epoch:** 0.84 ± 0.13 **DL (s):** 2.4 ± 3.5 **FPR/h:** 48.2 ± 33 **Norm. FPR:** 105.1 ± 72	**Sens/sz:** 0.18 **Sens/epoch:** 0.16 ± 0.34 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 1.5 ± 1.5 **FPR/h:** 5.2 ± 7.2 **Norm. FPR:** 99.3 ± 100.9	1
MUN3	**Sens/sz:** 1.00 **Sens/epoch:** 0.74 ± 0.22 **Spec/epoch:** 0.81 ± 0.14 **DL (s):** 1.5 ± 2.3 **FPR/h:** 53.4 ± 35.4	**Sens/sz:** 0.83 **Sens/epoch:** 0.69 ± 0.35 **Spec/epoch:** 0.83 ± 0.13 **DL (s):** 1.8 ± 3.6 **FPR/h:** 44.8 ± 29.1	**Sens/sz:** 0.83 **Sens/epoch:** 0.34 ± 0.23 **Spec/epoch:** 0.97 ± 0.04 **DL (s):** 7.8 ± 8.4 **FPR/h:** 12.5 ± 12.7	1
TUH1	**Sens/sz:** 1.00 **Sens/epoch:** 0.83 ± 0.17 **Spec/epoch:** 0.92 ± 0.14 **DL (s):** 0.0 ± 0.0	**Sens/sz:** 1.00 **Sens/epoch:** 0.95 ± 0.05 **Spec/epoch:** 1.0 ± 0.0 **DL (s):** 1.7 ± 3.2	**Sens/sz:** 1.00 **Sens/epoch:** 0.79 ± 0.16 **Spec/epoch:** 1.0 ± 0.0 **DL (s):** 0.0 ± 0.0	3
TUH2	**Sens/sz:** 1.00 **Sens/epoch:** 1.0 ± 0.0 **Spec/epoch:** 0.89 ± 0.09 **DL (s):** 0.0 ± 0.0	**Sens/sz:** 0.8 **Sens/epoch:** 0.77 ± 0.4 **Spec/epoch:** 0.66 ± 0.22 **DL (s):** 0.34 ± 0.99	**Sens/sz:** 0.8 **Sens/epoch:** 0.77 ± 0.39 **Spec/epoch:** 0.96 ± 0.04 **DL (s):** 0.3 ± 0.04	1
TUH3	**Sens/sz:** 0.91 **Sens/epoch:** 0.91 ± 0.29 **Spec/epoch:** 0.78 ± 0.28 **DL (s):** 0.0 ± 0	**Sens/sz:** 0.91 **Sens/epoch:** 0.91 ± 0.29 **Spec/epoch:** 0.86 ± 0.15 **DL (s):** 0.0 ± 0.0	**Sens/sz:** 0.91 **Sens/epoch:** 0.91 ± 0.29 **Spec/epoch:** 0.86 ± 0.28 **DL (s):** 0.0 ± 0.0	1
TUH4	**Sens/sz:** 0.9 **Sens/epoch:** 0.73 ± 0.3 **Spec/epoch:** 0.87 ± 0.22 **DL (s):** 2.0 ± 2.8	**Sens/sz:** 0.8 **Sens/epoch:** 0.56 ± 0.33 **Spec/epoch:** 0.91 ± 0.17 **DL (s):** 1.8 ± 2.9	**Sens/sz:** 0.6 **Sens/epoch:** 0.36 ± 0.37 **Spec/epoch:** 1.00 ± 0.00 **DL (s):** 2.5 ± 3.2	2
TUH5	**Sens/sz:** 1.00 **Sens/epoch:** 0.97 ± 0.06 **Spec/epoch:** 0.97 ± 0.05 **DL (s):** 0 ± 0	**Sens/sz:** 0.88 **Sens/epoch:** 0.79 ± 0.32 **Spec/epoch:** 0.89 ± 0.15 **DL (s):** 0 ± 0	**Sens/sz:** 0.88 **Sens/epoch:** 0.77 ± 0.3 **Spec/epoch:** 0.99 ± 0.02 **DL (s):** 0 ± 0	2

which has very high sensitivity and low detection delay, showing potential for injury prevention. Patient TUH2 shows the best sensitivity with EEG analysis, although ECG has similar results with reduced false positives, and is a less intrusive sensor and with less noise on ambulatory. Patient TUH3 has very high sensitivity per seizure in EEG, and an increased specificity in the multimodal approach, although with this data fusion methodology the sensitivity slightly decreases. As such, this patient would benefit from having a wearable with few EEG electrodes. TUH4 benefits from having a multimodal approach in terms of decrease of false positives, sacrificing sensitivity per epoch. Finally, TUH5 has a decrease in sensitivity per seizure with the multimodal approach, but an increase in specificity per epoch.

MUN1 shows similar behavior to TUH4, but benefits from the multimodal approach in terms of false positives. In MUN2, there is a decrease in the sensitivity per seizure in the multimodal approach, however, the decrease in false positives compensates it, as the unimodal solutions have too many false positives per hour. Finally, in MUN3 the multimodal approach shows higher specificity (and lower false positive rate) but lower sensitivity, and also a lower delay when compared to EEG.

Regarding normalized FPR, false positives seem most penalizing for MUN2 and MUN3, as this metric achieves higher values compared to MUN1, for both unimodal and multimodal approaches.

5.4.2 Discussion

For both MUN1 and TUH4, a high sensitivity for seizure detection using ECG was likely to happen, as temporal lobe seizures, as in 82% of the patients with epilepsy, seizures are accompanied by ictal sinus tachycardia (45), due to the activation of the central autonomic network caused by epileptic discharges. This occurs especially in patients with generalized tonic-clonic seizures and focal impaired awareness seizures, originating from the temporal lobe (46). However, it is not clear if the seizures were detected due to tachycardia or to other factors.

Data from University of Munich is more representative of realistic conditions, regarding the fact that it is composed of long term recordings, in which patients do their daily routine while in bed. On the contrary, TUH data has removed irrelevant signal segments that could be important for validating an alarm system, for a proper calculation of FPR/h. As such, the interictal to ictal ratio is lower than data from Munich, on average.

It is possible that the high number of false positives is derived from the occurrence of seizures leading to a high amount of movement artifacts on the signals during the ictal period. In the ictal period, movement occurs naturally, and could be misclassified as seizures. With this in mind, the algorithm might be more useful as a nocturnal seizure detector, as information about sleep stages can also give some information regarding the amount of movement.

These results show that logically combining the output of two signals with an "AND" operator is most beneficial when both signals give good results. If one seizure is not detected with one of the classifiers, the multimodal approach's sensitivity also decreases.

Results in the literature cannot be directly comparable with the obtained here, as the used dataset (and often, the performance metrics) is different. Despite these limitations, a comparison with the most related works in the state of the art (section 4.6) is presented here. Future researchers can use the same patients to validate their solutions and compare their results to the ones presented here.

Compared to the globality of the works presented in section 4.6, the obtained sensitivity seems higher, but the specificity seems lower. Note that most works do not detail whether their sensitivity is per seizure or per epoch.

DeCooman et al. (51) presented an average sensitivity of 77.6% and a FAR of 2.56/night, for a patient-specific seizure detector using ECG. Assuming 8 hours of sleep, the FAR is around 0.32/h. The obtained approach for ECG has an average sensitivity per seizure of 89%, for all 8 patients, but a much higher FAR for ECG (which was only evaluated in Munich patients).

Regarding the multimodal approaches, the closest work to this dissertation (122) presents a multimodal approach with an average sensitivity per seizures of 100% for 7 out of 10 patients, averaging 70%, using HR, EDA and SpO2, 60% personalizing those parameters, for 10 other

patients, and 67% for three other patients, using EEG with the previous sensors. In this work, the average sensitivity per seizure is of 75%, surpassing that work. No FAR/h is reported.

Other works which use different sensors for multimodal, patient-independent classification (120; 121; 93), achieve higher sensitivities and lower FAR/h (presented in the state of the art), but are hard to compare as the methodology is very different.

5.5 Pipeline parameter analysis

Next, different parameters of the classification pipeline (data augmentation, resampling, class imbalance) are varied, and its influence on classification is reported. Due to computation time and to facilitate reading this dissertation, only a few examples are given, from Munich patients, which were the hardest to classify according to the previously presented results.

Moreover, there is a focus on providing understandable metrics to aid the healthcare professional's decision making in optimizing the metrics for each patient, such as displaying plots which can be placed in an interface.

5.5.1 Effect of data augmentation

The impact of data augmentation in the dataset was studied on the two patients with the least amount of seizures (MUN2 and MUN3). The goal was to identify the impact of having substantial data transformations on the performance on the validation set. A leave one seizure out approach was performed and the sensitivity per epoch and specificity were calculated and averaged for all folds, as these metrics vary a lot with these transformations. No post processing was applied.

After iteratively finding the optimal range of values for data augmentation that allowed for a faster model convergence, by subjective analysis, the following data augmentation schemes, whose results are present in Fig 5.7, were attempted:

1. **None:** Raw data with no augmentation.

2. **Small:** Random noise with a standard deviation of 0.05, and random scaling with a normal distribution having standard deviation of 0.01 was applied (as in section 5.3)

3. **Jit0.30.60.9:** Random noise with a standard deviation of 0.3, 0.6 and 0.9 multiplied by the average of the absolute value of the signal (normalization according to signal amplitude)

4. **Jit0.30.60.9Scal0.05:** Random noise with a standard deviation of 0.3, 0.6 and 0.9 and random scaling with a standard deviation of 0.05, both multiplied by the average of the absolute value of the signal

5. **Jit0.3:** Random noise with a standard deviation of 0.3, scaling with a standard deviation of 0.05, multiplied by the average of the absolute value of the signal.

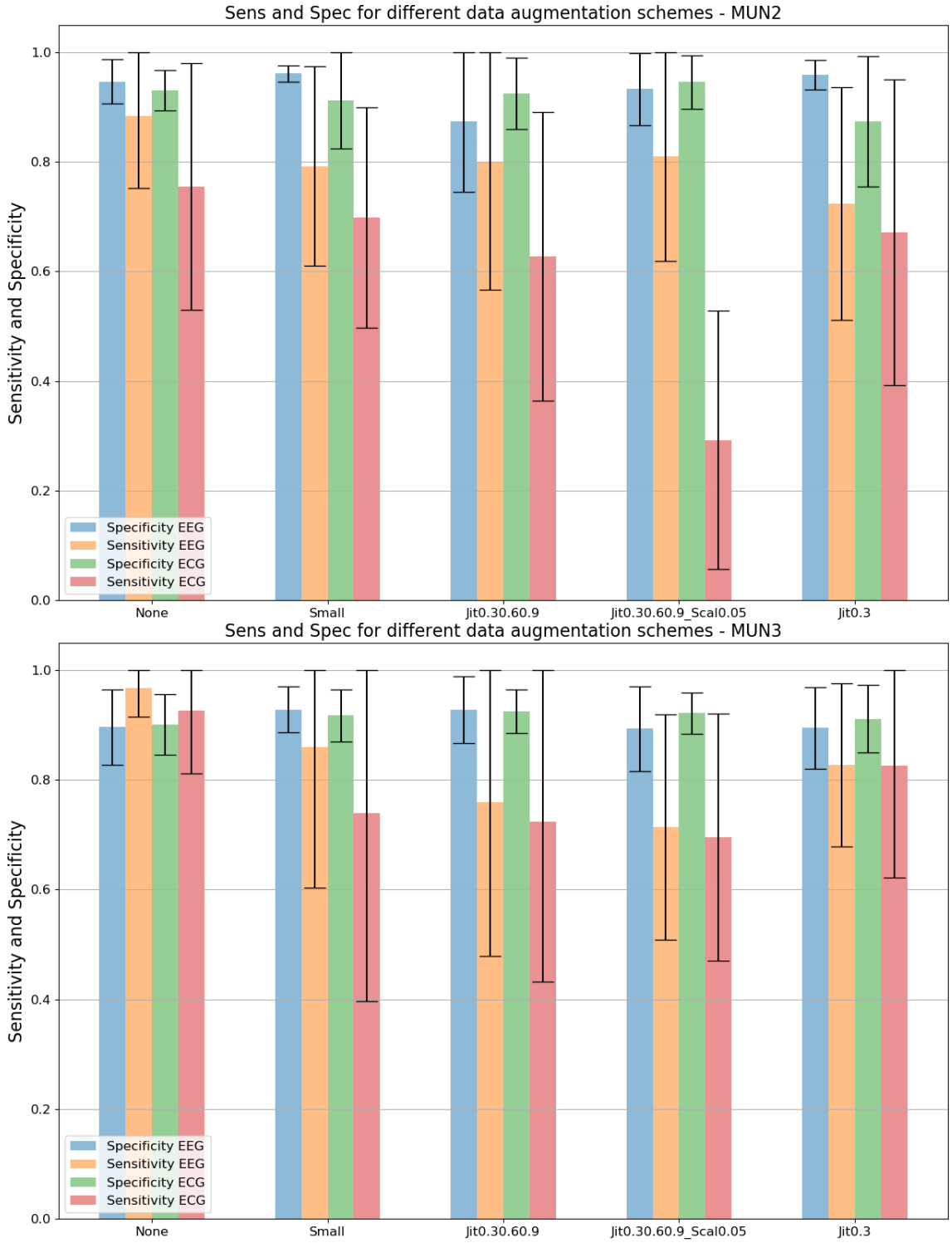


Figure 5.7: Performance for the different types of data augmentation applied in MUN2 and MUN3.

Due to the differences in frequency and amplitude of EEG and ECG, the performance of data augmentation will be analyzed individually for each signal.

First, for EEG, the small amplitude data augmentation (*Small*) led to an increase in specificity (and decrease in its standard deviation) when compared to no data augmentation. On the other hand, sensitivity slightly decreased and its standard deviation increased. This behavior was seen for both MUN2 and MUN3. *Jit0.3* behaved similarly.

In *Jit0.30.60.9*, there is a decrease in performance for both patients compared to no data augmentation, except in MUN1, where specificity increased. For *Jit0.30.60.9Scal0.05*, a similar behavior was observed.

For all cases, standard deviation in regards to sensitivity increased by more than twice when compared to the value with no data augmentation. On the other hand, specificity increased, on most cases, which leads to a decrease in the number of false alarms in the seizure detector. This can mean that the data augmentation strategies used for EEG are better for the interictal period than for the ictal period. For the ictal period, it is possible that the transformations lead to the data being too different from the validation set.

On the other hand, for ECG, in MUN2, all types of data augmentation lead to an increase in the standard deviation for sensitivity and specificity compared to no augmentation, and to a decrease in performance. However, for MUN3, despite the decrease in sensitivity and an increase in its standard deviation, specificity is better for all types of data augmentation. An hypothesis to this is that transformations to ECG applied during the ictal period may lead to more difficulties in estimating slight variations in heart rate or signal shape.

The literature shows that the effect of data augmentation on time series classification lacks a thorough study (128). This study has demonstrated that in multimodal approaches, data augmentation should be biosignal-dependent, and perhaps class-dependent when the classes are too different.

5.5.2 Effect of class imbalance on classifier performance

To assess the influence of class imbalance during training on the performance of the classifier, the two patients with least amount of seizures from the Munich dataset (MUN2 and MUN3) were used. After the initially tested interictal to ictal subsampling ratio of 8 was presented, a detailed analysis is shown here, by varying this ratio to 1, 5, 10, 15 and 20.

The specificity, sensitivity per seizure and sensitivity per epoch were calculated with no post processing, to determine the influence of interictal to ictal ratio on the raw output. Detection delay is not plotted as the results were shown not to vary significantly. Figure 5.8 reports the results for MUN2 and MUN3.

For MUN2, specificity decreases by around 30% when the resampling ratio is 1, compared to the remaining ratios. With the increase in the resampling ratio, specificity for EEG and ECG is stable, and variations are within the expected range of model stochasticity. For EEG, sensitivity is also kept within the same range, while for ECG it is more variable and with higher standard

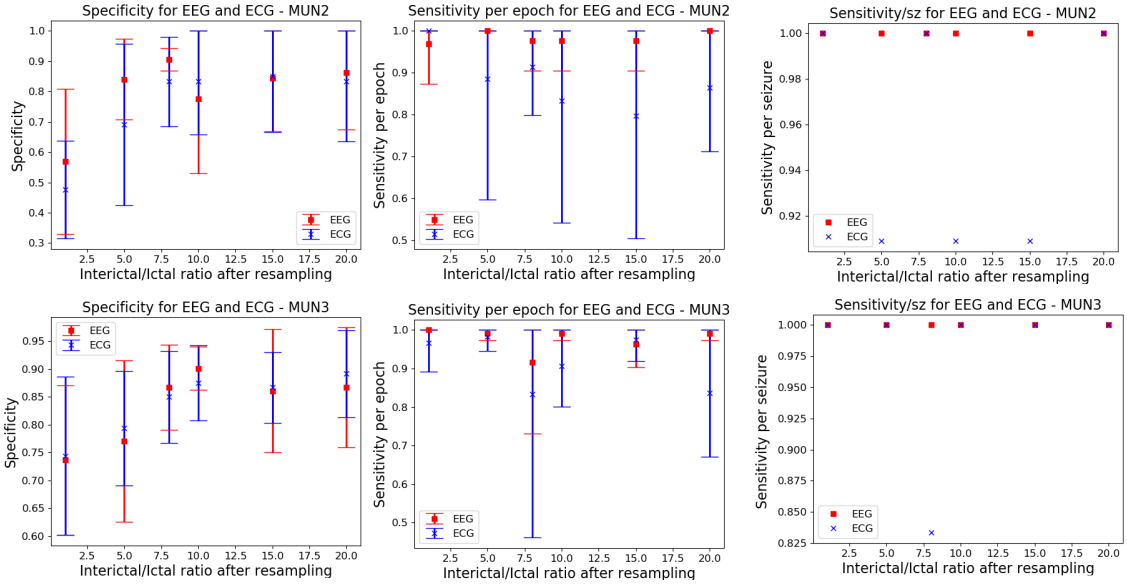


Figure 5.8: Effect of resampling the interictal samples with an interictal/ictal ratio of 1, 5, 8, 10, 15 or 20 in model performance.

deviation than ECG. Sensitivity per seizure does not have a visible pattern, with variations between 92% and 100%.

For MUN3, the decrease in specificity for the resampling ratio of 1 is of around 10%. Similarly to MUN2, specificity does not show much variability between the ratio of 8 and 20, and there is not a clear pattern regarding the changes in sensitivity per epoch or per seizure.

It can be concluded, that overall, the model responds well to class imbalance, and can maintain its performance with slight increases in class imbalance. For lower resampling ratios, the amount of data is also smaller, leading to a worse performance - as such, there is a trade off between the amount of data and "resistance" to class imbalance, which appears to hit a plateau starting at a given ratio.

5.5.3 Solving class imbalance - Focal Loss

As another alternative to diminishing the impact of class imbalance, instead of applying class weights to the classification, focal loss as a loss function was used, for all long term monitoring patients (Munich).

Focal Loss (Eq. 5.1) is a cross-entropy loss, weighting the contribution of each sample to the loss depending on the classification error. This way, easily classified samples have less impact in the loss, which can implicitly solve class imbalance by focusing on problematic classes.

$$FL = -\sum_{i=1}^{C=2} (1 - s_i)^{\gamma} t_i log(s_i), \tag{5.1}$$

where γ is modulating factor to reduce the influence of correctly classified samples in the loss and s_i is the classifier score for each class i in C.

Table 5.6 shows the obtained results for classification using the LSTM network with focal loss.

Table 5.6: Summary of the results for EEG, ECG and multimodal (MM) fusion (LSTM), using the focal loss.

Patient	EEG	ECG	MM	Alarm size
MUN1	**Sens/sz:** 1.00 **Sens/epoch:** 0.94 ± 0.09 **Spec/epoch:** 0.98 ± 0.01 **DL (s):** 1.4 ± 3.6 **FPR/h:** 3.65 ± 3.96	**Sens/sz:** 0.95 **Sens/epoch:** 0.62 ± 0.26 **Spec/epoch:** 0.95 ± 0.06 **DL (s):** 4.5 ± 6.3 **FPR/h:** 12.4 ± 14.5	**Sens/sz:** 0.95 **Sens/epoch:** 0.59 ± 0.24 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 6.0 ± 7.1 **FPR/h:** 1.0 ± 1.3	2
MUN2	**Sens/sz:** 0.91 **Sens/epoch:** 0.89 ± 0.29 **Spec/epoch:** 0.91 ± 0.07 **DL (s):** 0.3 ± 0.9 **FPR/h:** 28.6 ± 18.9	**Sens/sz:** 0.82 **Sens/epoch:**0.7 ± 0.37 **Spec/epoch:** 0.83 ± 0.15 **DL (s):** 1.7 ± 2.05 **FPR/h:** 34.9 ± 31.3	**Sens/sz:** 0.72 **Sens/epoch:** 0.6 ± 0.4 **Spec/epoch:** 0.98 ± 0.02 **DL (s):** 1.9 ± 2.1 **FPR/h:** 6.2 ± 7.5	1
MUN3	**Sens/sz:** 1.00 **Sens/epoch:** 0.87 ± 0.16 **Spec/epoch:** 0.96 ± 0.03 **DL (s):** 1.5 ± 2.3 **FPR/h:** 9.2 ± 6.7	**Sens/sz:** 0.83 **Sens/epoch:** 0.62 ± 0.36 **Spec/epoch:** 0.93 ± 0.05 **DL (s):** 5.4 ± 11 **FPR/h:** 15.4 ± 9.8	**Sens/sz:** 0.83 **Sens/epoch:** 0.59 ± 0.34 **Spec/epoch:** 0.99 ± 0.02 **DL (s):** 6.0 ± 10 **FPR/h:** 3.6 ± 3.7	3

Although focal loss can effectively lead to some correction of class imbalance, the presented performance is lower than binary cross-entropy with class weights compared to section 5.4.1, as the number of false positives seems to increase, while the overall sensitivity decreases.

5.5.4 Optimizing threshold

In section 5.4.1, the shown results assume a threshold of 0.5 applied on the probabilities output by the classifier. However, this value can be optimized for obtaining the best trade off between sensitivity and specificity, through the ROC curve. Since cross validation (leave one seizure out) was performed, a number of ROC curves equal to the number of folds can be generated. An example of the ROC curve generated is shown on Figure 5.9.

In cases of classification with class imbalance, the optimal threshold for classification, OptThresh, can be found by seeking the value i which maximizes the harmonic mean on the ROC curve for each individual biosignal (129), as defined by Equation 5.2:

$$OptThresh = max_i(\frac{2 * sensitivity(i) * specificity(i)}{sensitivity(i) + specificity(i)}) \tag{5.2}$$

The average of all optimal thresholds was found. Points which were outside the range μ(OptThresh) $\pm$ σ (OptThresh) were considered to be outliers and were therefore removed. Although this could lead to a small increase in overfitting, as a reduced number of points are being removed, it led to an improval in cross validation results.

Table 5.7 shows the obtained performance for the automatically found threshold, keeping the same amount of post processing as previously.

Compared to the previous results, MUN1 shows an increase in sensitivity per epoch at the cost of a slight increase in FPR/h. MUN2 has decrease in its performance, in terms of sensitivity and

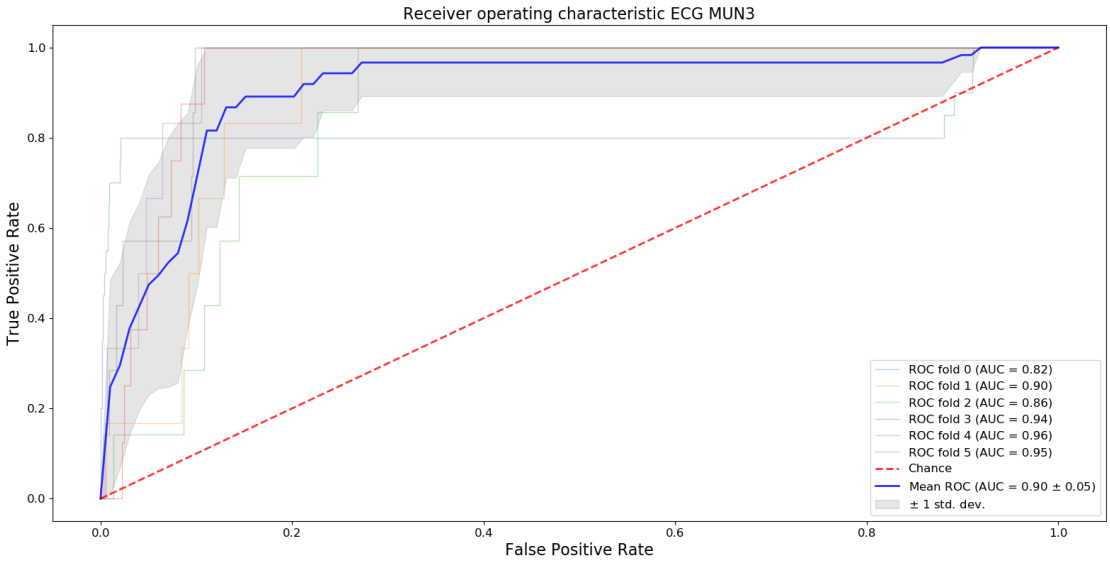


Figure 5.9: Cross-validation ROC curve for EEG, for MUN3.

false alarm rate. MUN3 has an increase in the number of detected seizures, with a slight increase in false positive rate.

From the presented results, it is shown that an automatic selection of the threshold can lead to significant differences in terms of performance. It is not clear if this methodology leads to an improvement, as the number of patients is very low. However, for some cases, this may be beneficial, as an increased sensitivity per epoch might lead to the possibility of increasing the alarm size, reducing the amount of false positives. For instance, for MUN3, an alarm size of 5 leads to detecting 85% of all seizures, with a FPR of 0.3 ± 1.09. With the threshold fixed at 0.5, the FPR/h is changed to 0 but the number of detected seizures declines to 58%.

5.5.5 Effect of alarm size

Although in the previous section the alarm size was determined iteratively, it is important to know for a specific patient how the sensitivity and specificity is altered by an increase in an alarm size. All this information merged together may assist in determining the best trade off between sensitivity and specificity, since missing the detection of seizures with short duration (such as 6 seconds, equivalent to 2 epochs) is not as critical as missing longer seizures, which are more dangerous and may lead to more severe injuries.

Figure 5.10 presents the seizure's true length as a function of alarm size, also showing which seizures were and were not detected (triangles and circles, respectively), for patient MUN3.

It can be seen that for ECG, the longest seizure is not detected even with alarm size 1 epoch. For an alarm size of 8 epochs, no seizure is detected. For EEG, all seizures are detected until an alarm size of 4 epochs. However, the multimodal approach cannot detect the longest seizure, since

Table 5.7: Summary of the results for EEG, ECG and multimodal fusion (LSTM), using the automatically found threshold.

Patient	EEG	ECG	MM	Alarm size
MUN1	**Sens/sz:** 1.00 **Sens/epoch:** 0.92 ± 0.12 **Spec/epoch:** 0.98 ± 0.02 **DL (s):** 1.7 ± 3.9 **FPR/h:** 4.4 ± 3.6	**Sens/sz:** 1.00 **Sens/epoch:** 0.94 ± 0.08 **Spec/epoch:** 0.79 ± 0.22 **DL (s):** 0.47 ± 1.46 **FPR/h:** 15.8 ± 16.2	**Sens/sz:** 1.00 **Sens/epoch:**0.86 ± 0.12 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 2.4 ± 4.1 **FPR/h:** 3.3 ± 2.6	2
MUN2	**Sens/sz:** 1.00 **Sens/epoch:** 0.86 ± 0.20 **Spec/epoch:** 0.88 ± 0.11 **DL (s):** 1.1 ± 1.9 **FPR/h:** 29.9 ± 25.9	**Sens/sz:** 1.00 **Sens/epoch:**0.92 ± 0.125 **Spec/epoch:** 0.61 ± 0.301 **DL (s):** 0.81 ± 1.34 **FPR/h:** 54.6 ± 31.3	**Sens/sz:** 0.81 **Sens/epoch:** 0.69 ± 0.63 **Spec/epoch:** 0.96 ± 0.06 **DL (s):** 1.6 ± 2.1 **FPR/h:** 14.1 ± 19.3	1
MUN3	**Sens/sz:** 1.00 **Sens/epoch:** 0.84 ± 0.09 **Spec/epoch:** 0.96 ± 0.04 **DL (s):** 3.5 ± 4.4 **FPR/h:** 10.6 ± 11.15	**Sens/sz:** 1.00 **Sens/epoch:** 0.76 ± 0.21 **Spec/epoch:** 0.93 ± 0.03 **DL (s):** 5.5 ± 8.9 **FPR/h:** 17.4 ± 8.2	**Sens/sz:** 1.00 **Sens/epoch:** 0.69 ± 0.2 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 6.0 ± 8.7 **FPR/h:** 3.6 ± 3.9	3

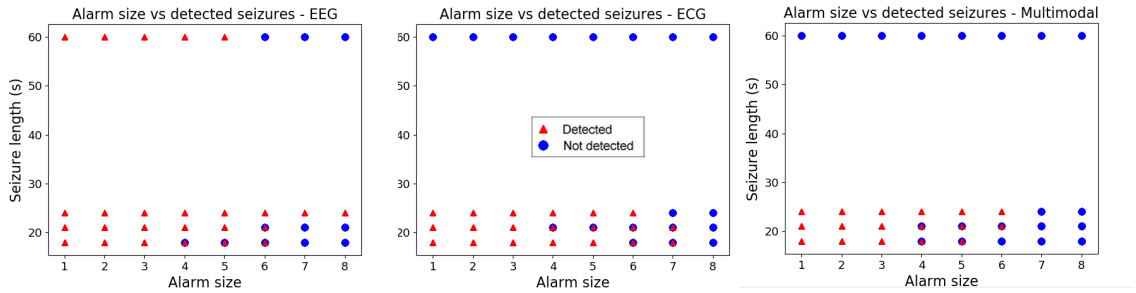


Figure 5.10: Effect of alarm size (epochs) on the number of detected seizures for ECG, EMG and Multimodal. Triangles show detected seizures while circles show non detected seizures.

a logical AND operation was performed between EEG and ECG, showing one of the limitations of this combination.

Figure 5.11 show the sensitivity per epoch and the specificity plotted against the alarm size.

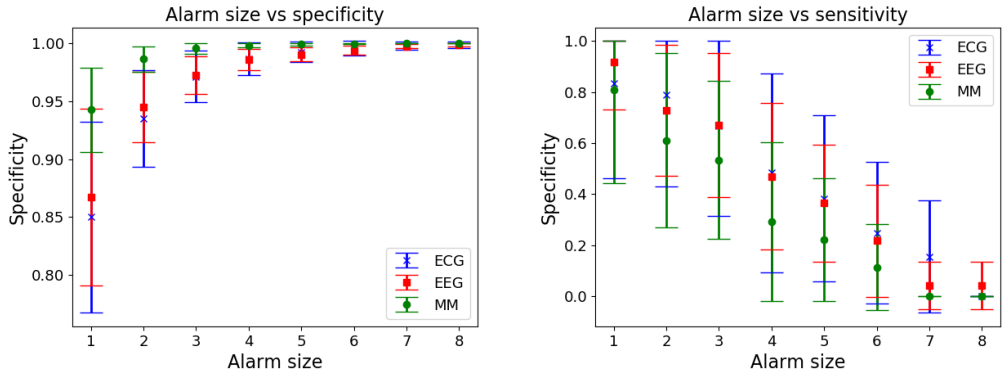


Figure 5.11: Effect of alarm size (epochs) on the number of detected seizures for ECG, EMG and Multimodal.

There is an exponential decrease in sensitivity with the increase in alarm size, and as previously

seen, the sensitivity of the multimodal approach is lower than unimodal. On the other hand, specificity exponentially increases with the increase in alarm size, with the multimodal approach reaching the plateau of around 1 quicker than the unimodal approaches. Also, it is visible that the standard deviation also decreases. Although only a single example is shown here, this behavior is common to all patients.

5.5.6 Effect of number of training samples

The number of seizures each patient has on the EMU is very variable, and can sometimes be very low (1-3) per patient. As such, it is important to assess the capability of the classifier to learn how to correctly detect the patient's seizures with little data, while keeping the amount of false positives to a minimum.

For all Munich patients, the number of test seizures was iteratively increased. Seizures were randomly shuffled and the experiment was replicated 5 times, to have some variability in the test seizures while reducing the computational time needed. Sensitivity per seizure, specificity and sensitivity per epoch were calculated with no post processing, for EEG, ECG and a multimodal approach (combining both predictions using a logical "AND").

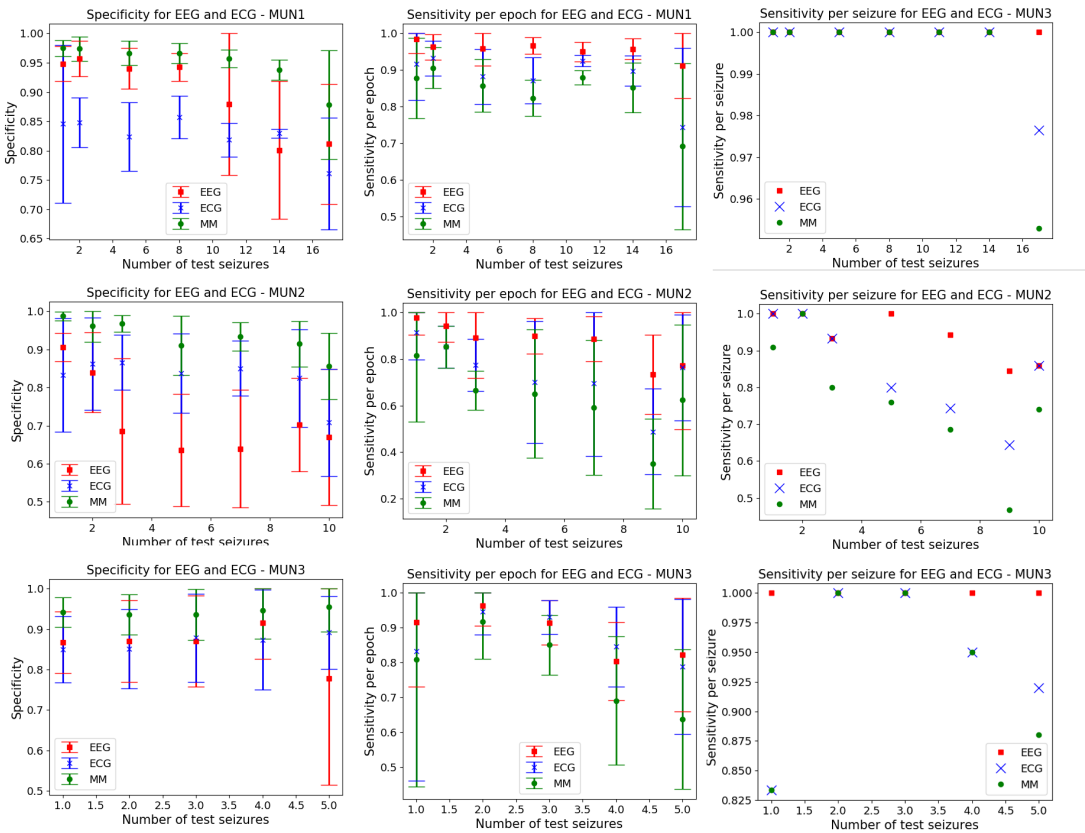


Figure 5.12: Effect on the number of training samples on classifier's performance, for all long term monitoring recordings.

For MUN1, it can be seen that all seizures are detected, for EEG, ECG and MM, except for when there's a single training sample. However, specificity gradually decreases with the decrease in the number of training samples, for unimodal approaches. For the multimodal approach, specificity is more stable with the reduction in training samples. Sensitivity per epoch is maintained mostly for EEG, but shows a slightly higher variation for ECG. By combining both signals, sensitivity per epoch is always reduced, as expected.

For MUN2, sensitivity per seizure shows a trend of decrease with an increase in the number of test seizures. Specificity for EEG decays starting on 3 test seizures, while for ECG it is more stable. Lastly, sensitivity per epoch is stable for EEG while there is a decreasing trend in ECG and multimodal approaches.

For MUN3, there is a decay in the performance in sensitivity per seizure seen in ECG and multimodal approaches. With an increase in the number of training seizures, an increase in performance is shown, since for the case of one test seizures, a leave one seizure out approach was used, instead of randomly picking 5 seizures. Specificity and sensitivity do not show significant variations. Even with just a single test seizure, sensitivity per seizure is above 87%.

According to the results presented above, LSTM model trained with class weights has good performance for seizure detection even with just a few training seizures. If using an unimodal approach, specificity may decay too much. However, these approaches show that by combining EEG and ECG we can significantly increase specificity, sacrificing slightly the performance for sensitivity.

5.6 Deep Learning Interpretability

The output of a machine learning algorithm is sometimes very hard to explain, as the number of dimensions is proportional to the number of features. Moreover, deep learning models are often *black boxes* as it is difficult to understand the transformations the data is being subjected to in the model's hidden layers. As such, healthcare professionals may be reluctant to accept new AI-based solutions. Researchers have been working on ways of providing interpretable metrics in cooperation with healthcare professionals (130).

Class Activation Map (CAM) is a way of providing an interpretable feedback which highlights the reason for a certain classification, and can be applied to both images and time series (125). This is only possible in networks that have a Global Average Pooling (GAP) Layer before the softmax/sigmoid classifier, such as the initially studied FCN and LSTM networks.

For a certain time series, An(x) represents the activation of a filter n on the last convolutional layer, at time x. For filter n, the output of the GAP layer is represented by Eq. 5.3.

$$f_n = \sum_n A_n(x) \tag{5.3}$$

w_n^c is the weight of the last softmax/sigmoid function for the output, from filter n and class c. As such, the input of the final softmax/sigmoid function, g_c, is given by Eq. 5.4.

$$g_c = \sum_n w_n^c \sum_x A_n(x) \tag{5.4}$$

The class activation map for class c is then given by Eq. 5.5.

$$M_c = \sum_n w_n^c A_n(x) \tag{5.5}$$

The CAM indicates how important is each activation at a location x that leads to the classification of a given time series as ictal, preictal or interictal.

Figure 5.13 shows an example of an ECG recording in the interictal period (left) and ictal period (right). For the normal ECG, the algorithm gives preference to the ECG peaks, perhaps indicating that the natural periodicity of the signal is crucial for characterizing it as interictal. On the ictal segment, the classifier gives more importance for the abrupt changes in the signal (artifacts caused by seizure movement), as this is a violent tonic clonic seizure.

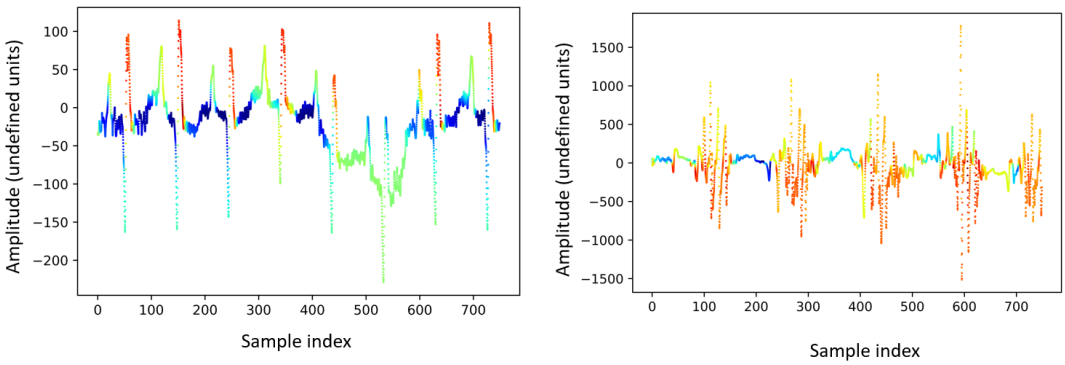


Figure 5.13: Class Activation Map applied to ECG, for the interictal period (left) and ictal period (right).

For signals with more subtle changes, this tool may be useful for helping gain knowledge on the properties of signals which are altered during seizures, and making deep learning models less of a black box.

Chapter 6

Optimization for a wearable device

This chapter contains preliminary work on developing a prototype of the obtained solution for a wearable device.

6.1 Embedded AI computing Single Board System

Recently, very complex algorithms for seizure detection have been implemented on small devices such as IBM's TrueNorth neuromorphic chip (131). This enables the creation of small, implantable devices, with low power consumption, but with the computing capacity of a supercomputer - such devices are named embedded AI computing single board systems.

Besides TrueNorth's chip, NVIDIA Jetson TX1 (132) includes the latest technology for deep learning and Graphics Processing Unit (GPU) computing, making it ideal for embedded Artificial Intelligence computing. Jetson's technical specifications are shown in table 6.1.

NVIDIA Jetson TX1 was used for initiating the prototype of this solution implemented on a portable system. This device is recommended to be used only for inference (i.e., real-time prediction). Training is expected to be performed in a host computer or a cloud instance. Regular Python code can be run on Jetson, using pre-installed libraries, as Jetson uses Linux as its Operating System.

After installing NVIDIA's drivers, Python and the typical libraries for machine learning (Keras, Scikit-learn) Jetson was ready to be used.

Table 6.1: NVIDIA Jetson TX1 Technical Specifications

Device	Brand
GPU	256-core NVIDIA Maxwell™ GPU
CPU	Quad-Core ARM® 64bits Cortex®-A57 MPCore
Memory	4GB 64-bit LPDDR4 Memory
Storage	16GB eMMC 5.1

6.1.1 Real time classification - Simulation

The viability of running the developed architecture in NVIDIA Jetson was evaluated by running a simulated system for data transmission and aggregation, described in this subsection.

From a laptop computer, two separate Python processes were ran. This simulated two different wearable devices - each process transmitted data from TUH1 subject to a network which received samples from ECG and EEG, interleaved, at a frequency of 256Hz. The samples were then pulled by an NVIDIA Jetson device running a code which aggregated data points into 3 second epochs, and classified them using the pre-trained LSTM models for EEG and ECG. The classification output and the reconstructed signals are plotted in an interface on another computer/screen. Then, an alarm is triggered on such interface if there are N subsequent detections from both devices. Figure 6.1 shows a sequence diagram for the developed system.

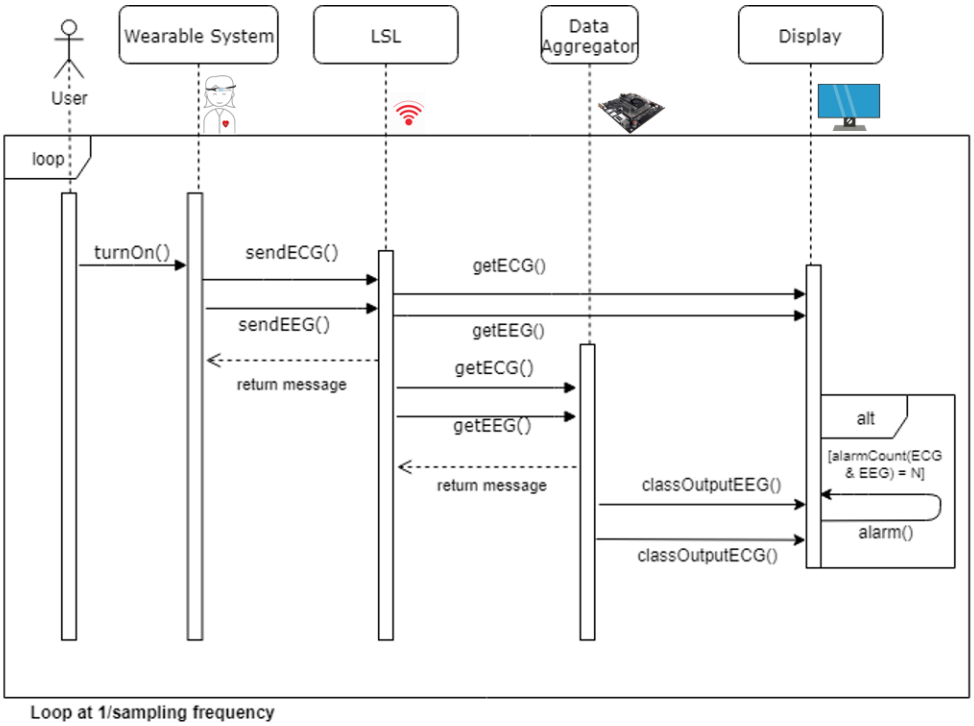


Figure 6.1: Sequence diagram for the near real time classification system.

For data transmission, the Lab Streaming Layer (LSL) system was used. LSL is an open source system developed by Swartz Center for Computational Neuroscience [1] , that allows exchange of time series between devices in real time, and recording from various devices. Pylsl [2] library, a Python wrapper for the C++ library liblsl, was specifically compiled for running in ARM64 processor (such as Jetson) and used for establishing this connection.

[1] SCCN - https://sccn.ucsd.edu/
[2] Pylsl - https://github.com/chkothe/pylsl

LSL works by sending a stream outlet using the local network clock, a container to which we push our data through. By creating an outlet, the stream is made visible to a collection of devices (defined by the network settings), that can be read using an inlet which detects a network of a specific type (or ID). Data is transmitted through Transmission Control Protocol (TCP).

As the offline results were already calculated in section 5.4.1 for patient TUH1, code execution time was calculated using Jetson's GPU. For program boot and model loading, it took on average 12.9 seconds, while an average of 50 classifications took 0.69 ± 0.17 s. Summing this with the 6s needed for generating an epoch (as reception of EEG and ECG samples is interleaved), the total detection delay is 6.69 ± 0.17 summed with the detection delay of each patient.

Overall, run time is very low. With sample overlap, which was not performed, this initial delay of epoch generation can also be decreased. These results show the potential of having a smaller system with computing capabilities similar to Jetson TX1 implemented on a device which processes data acquired from wearable devices.

6.1.2 Real time classification - Wearable devices

After validating the dataflow of the simulation, two wearables were used for further developing this prototype: Muse 2: Brain Sensing Headset and OpenBCI's ECG electrode, which transmit data at 250Hz.

Muse [3] is a headband which measures EEG, heart rate (PPG+Pulse Oximetry), Movement (Accelerometer) and Breathing (PPG+Gyroscope). It contains two electrodes on the forehead (AF8 and AF7), 2 behind the ears (TP9 and TP10) and 3 reference sensors. Data transmission to a computer is performed by Bluetooth Low Energy. For transmitting data with the same architecture as previously displayed, the muselsl Python toolbox [4] was used on Linux.

OpenBCI Cython Board [5] is an open source biosensing tool containing biopotential input channels (EEG, EMG and ECG), an accelerometer, Bluetooth Low Energy Communication. Data transmission by LSL was performed using the official OpenBCI-LSL Python toolbox [6].

The wearable acquisition setup is visible in Figure 6.2.

Correct data acquisition was validated by looking at the biosignals in real time. Afterwards, an interface was developed in Vispy [7] for plotting both the signals and the alarms in real time.

Vispy is a high-performance data visualization library for Python, which uses OpenGL for plotting signals in real time at a very high frequency. This is useful as matplotlib, the most common plotting library in Python, cannot handle high frequencies for displaying signals in real time. This can easily be transferred to a web application (through WebGL), turning this system multiplatform.

Vispy initially draws a blank screen, which is updated at the monitor's refresh rate. On each update, a chunk of samples is pulled, added at the end of an array, and plotted on the screen.

[3] Muse - https://choosemuse.com/
[4] MuseLSL - https://github.com/alexandrebarachant/muse-lsl.
[5] OpenBCI - https://openbci.com/
[6] OpenBCI LSL - https://github.com/OpenBCI/OpenBCI_LSL.
[7] Vispy - www.http://vispy.org/.

Figure 6.2: Wearable setup for data acquisition. On the forehead is the Muse device, and on the wrists are placed the two electrodes for measuring the ECG, connected to OpenBCI. The computer is planned to be used for data visualization. Jetson TX1 board is not shown.

The interface allows the user to zoom in all channels, amplifying the data's voltage, for a better visualization. It is not yet possible to do the same in the scale of time, something that more advanced data visualization softwares (e.g. those that come with VideoEEG devices) do.

The developed interface contains 7 rows, 4 for the EEG channels output from Muse, one for the ECG channel from OpenBCI and two for the classification results that NVIDIA Jetson TX1 sends to the computer running the interface, through LSL. If an alarm is triggered, the background turns red as a warning. An example is shown on Figure 6.3, using simulated data from TUH1 near a seizure.

The interface was also tested using the real time classification pipeline with Muse and OpenBCI. However, real time synchronized plotting of EEG and ECG from different devices was unsuccessful. The buffer extracted a different number of samples from each device in a single instant, and the interface was built using synchronous data streaming. As such, the data read from each sensor was a blocking operation, so a low number of frames per second was obtained.

Instead of having a chunk of points extracted in each loop of the program, a single point was extracted so the blocking read operation's duration was minimized. This led to a higher framerate, but the signal was too noisy, as filtering a single sample is not possible.

A video demonstration was recorded, and is available online [8]. Note that it is not possible to validate the solution's performance with these sensors, since the model is patient-specific.

[8] Demo - https://youtu.be/49Zt_AT3YXM

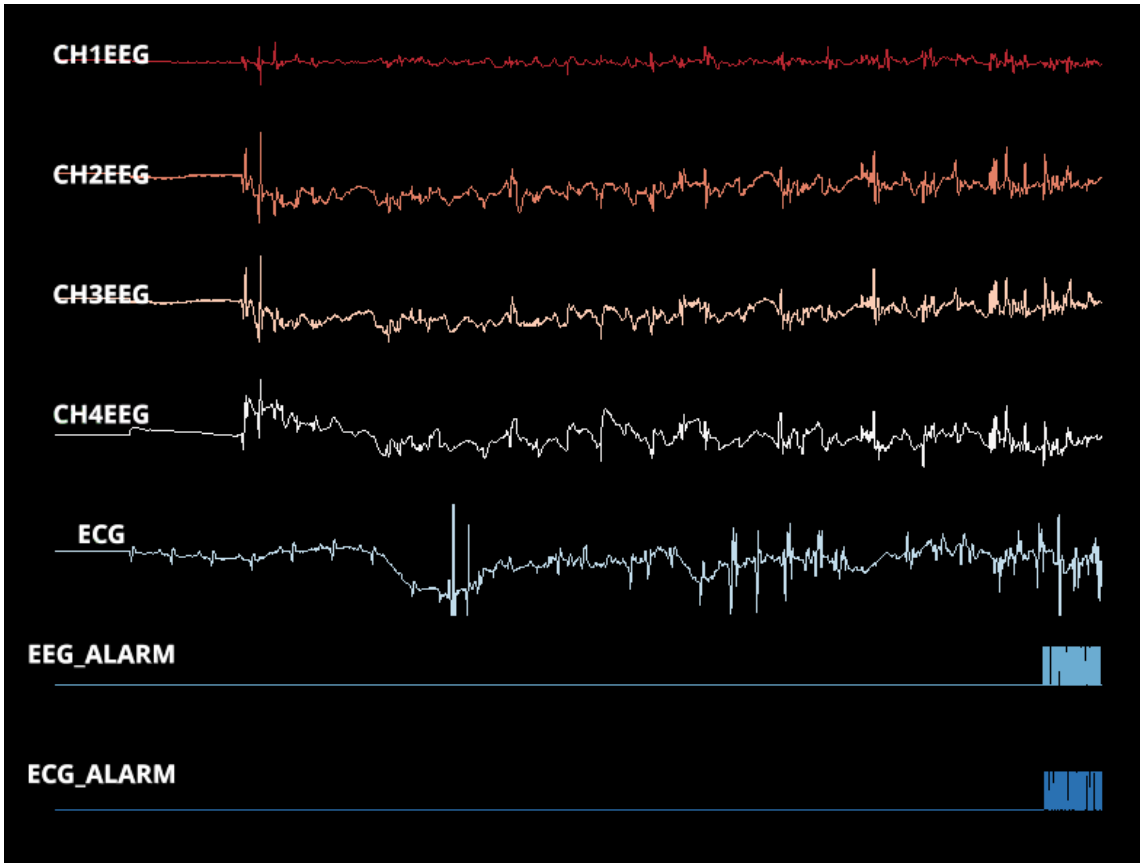


Figure 6.3: Developed interface using Vispy, running in real time using the buffered data.

Chapter 7

Epileptic Seizure Prediction

This chapter presents some initial results on epileptic seizure prediction, using all Munich patients (long term monitoring).

Since TUH patients have had segments of the signal cut, it does not make sense to try seizure prediction on that data, as it is not possible to know if the assumed preictal period is actually before the ictal.

These results are still very preliminary, as the main focus was on seizure detection. The goal of this chapter is to be an initial exploration for guiding future development of the project.

7.1 Classification

Previous studies have reported that the exact time when the preictal state starts is unknown, and may be patient dependent. A prediction time of minimum 10 seconds is considered to be adequate (86). As such, classification was modeled as a three-class problem, with the interictal and ictal period defined as in Chapter 5. However, the preictal period was considered as being 15, 30 and 45 seconds before the ictal period, in both training and validation samples. These values were picked to perform not a long term prediction, but an early detection, due to the high difficulty of finding predictive patterns with a low amount of data.

Undersampling the majority class was performed only in the interictal period, as the number of preictal samples is low when compared to interictal.

After some initial tests using the same parameters as chapter 5, prediction was found to give better results with the LSTM-based network having 150 units and the Time Distributed Dense layer with 100 layers. The categorical cross-entropy loss was used, since this is a multiclass problem. Similarly to Chapter 5, class weights during training were added according to the ratio between classes.

The best model was picked as the model having highest f1 score for class 1 (ictal) mutiplied by the f1 score for class 2 (preictal), so both ictal and preictal classification was maximized.

For post processing, similarly to the seizure detector presented in Chapter 5, epochs which had more than one preictal neighbor were considered to be preictal, and the same for ictal. Preictal

and ictal alarms were also triggered on N subsequent epochs with the same classification, where N is patient-dependent and optimized by cross-validation, to maximize the number of detectd and predicted seizures, except in selected cases.

Multimodal seizure prediction was obtained by using the logical AND operator for combining ictal predictions, and the logical OR for combining preictal, as it was seen that some seizures could only be predicted by one biosignal.

An example of the presented output, after post processing and multimodal fusion, is presented on figure 7.1.

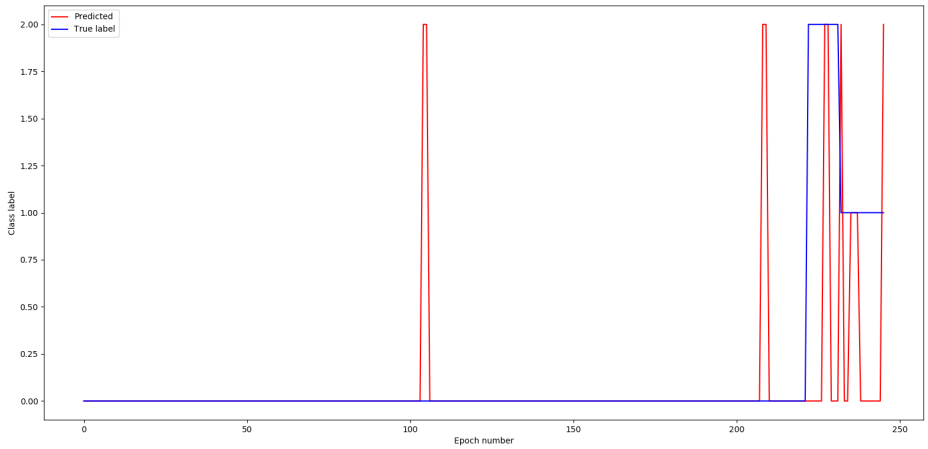


Figure 7.1: Example of the output for seizure prediction, showing a detected and predicted seizure with two false predictions.

7.2 Results

Classification results are presented for all Munich patients, in tables 7.1 to 7.3 (15, 30 and 45 seconds of assumed pre-ictal period).

Due to the high number of metrics presented, only the most relevant ones for having a minimally functioning system for clinical environment will be discussed: false positive rates, sensitivity per seizure state and prediction time.

For MUN1, an assumed pre-ictal period of 30 seconds gives the best results in terms of predicted seizures (74%), for the multimodal approach. The false positive rate for multimodal seizure detection (0.69 ± 1.25) is low, and all seizures are detected. For seizure prediction, FPR is higher than for 15 seconds (7.36 ± 5.35 vs 2.99 ± 5.53), but both are lower than for 45 seconds (15.31 ± 10.86). Prediction time is around 6-7 $\pm$ 5-6 seconds for all cases, which could mean that relevant changes for early seizure detection occur around this time.

Table 7.1: Summary of the results for EEG, ECG and multimodal fusion (LSTM), assuming 15 seconds of preictal time.

Patient	EEG	ECG	MM	Alarm size
MUN1	**Sens/epoch interictal:** 0.95 ± 0.04 *Prediction:* **Sens/sz:** 0.11 **Sens/epoch:** 0.02 ± 0.06 **DL (s):** -9.0 ± 6.0 **FPR/h:** 0.5 ± 0.78 *Detection:* **Sens/sz:** 1.00 **Sens/epoch:** 0.69 ± 0.23 **DL (s):** 6.6 ± 8.7 **FPR/h:** 10.8 ± 9.1	**Sens/epoch interictal:** 0.98 ± 0.03 *Prediction:* **Sens/sz:** 0.37 **Sens/epoch:** 0.2 ± 0.28 **DL (s):** -5.14 ± 3.48 **FPR/h:** 2.5 ± 5.51 *Detection:* **Sens/sz:** 1.00 **Sens/epoch:** 0.88 ± 0.13 **DL (s):** 2.53 ± 4.38 **FPR/h:** 3.62 ± 3.49	**Sens/epoch interictal:** 0.99 ± 0.02 *Prediction:* **Sens/sz:** 0.47 **Sens/epoch:** 0.22 ± 0.27 **DL (s):** -6.0 ± 4.47 **FPR/h:** 2.99 ± 5.53 *Detection:* **Sens/sz:** 1.00 **Sens/epoch:** 0.6 ± 0.21 **DL (s):** 7.58 ± 8.56 **FPR/h:** 1.09 ± 1.02	2
MUN2	**Sens/epoch interictal:** 0.63 ± 0.28 *Prediction:* **Sens/sz:** 0.73 **Sens/epoch:** 0.45 ± 0.33 **DL (s):** -4.5 ± 1.5 **FPR/h:** 75.66 ± 60.18 *Detection:* **Sens/sz:** 0.73 **Sens/epoch:** 0.53 ± 0.37 **DL (s):** 2.25 ± 1.3 **FPR/h:** 52.93 ± 50.9	**Sens/epoch interictal:** 0.69 ± 0.27 *Prediction:* **Sens/sz:** 1.00 **Sens/epoch:** 0.62 ± 0.22 **DL (s):** -3.82 ± 1.34 **FPR/h:** 73.58 ± 58.83 *Detection:* **Sens/sz:** 0.91 **Sens/epoch:** 0.7 ± 0.31 **DL (s):** 1.8 ± 1.99 **FPR/h:** 43.0 ± 56.25	**Sens/epoch interictal:** 0.55 ± 0.25 *Prediction:* **Sens/sz:** 1.00 **Sens/epoch:** 0.85 ± 0.09 **DL (s):** -3.55 ± 1.16 **FPR/h:** 84.71 ± 26.51 *Detection:* **Sens/sz:** 0.55 **Sens/epoch:** 0.36 ± 0.36 **DL (s):** 2.5 ± 2.06 **FPR/h:** 6.17 ± 13.9	1
MUN3	**Sens/epoch interictal:** 0.92 ± 0.05 *Prediction:* **Sens/sz:** 0.17 **Sens/epoch:** 0.13 ± 0.3 **DL (s):** -6.0 ± 0.0 **FPR/h:** 3.2 ± 3.09 *Detection:* **Sens/sz:** 1.00 **Sens/epoch:** 0.75 ± 0.26 **DL (s):** 5.5 ± 8.9 **FPR/h:** 18.85 ± 13.77	**Sens/epoch interictal:** 0.97 ± 0.02 *Prediction:* **Sens/sz:** 0.5 **Sens/epoch:** 0.4 ± 0.42 **DL(s):** - 5.0 ± 2.83 **FPR/h:** 1.33 ± 1.87 *Detection:* **Sens/sz:** 1.00 **Sens/epoch:** 0.71 ± 0.2 **DL(s):** 6.0 ± 5.74 **FPR/h:** 8.39 ± 5.0	**Sens/epoch interictal:** 0.98 ± 0.01 *Prediction:* **Sens/sz:** 0.5 **Sens/epoch:** 0.43 ± 0.45 **DL(s):** - 5.0 ± 2.83 **FPR/h:** 4.32 ± 2.74 *Detection:* **Sens/sz:** 0.83 **Sens/epoch:** 0.51 ± 0.31 **DL(s):** 8.4 ± 8.98 **FPR/h:** 2.36 ± 1.32	2

Table 7.2: Summary of the results for EEG, ECG and multimodal fusion (LSTM), assuming 30 seconds of preictal time.

Patient	EEG	ECG	MM	Alarm size
MUN1	**Sens/epoch interictal:** 0.94 ± 0.06 *Prediction:* **Sens/sz:** 0.37 **Sens/epoch:** 0.09 ± 0.14 **DL(s):** - 15.0 ± 9.49 **FPR/h:** 4.2 ± 3.26 *Detection:* **Sens/sz:** 1.0 **Sens/epoch:** 0.62 ± 0.16 **DL(s):** 6.0 ± 5.51 **FPR/h:** 10.78 ± 11.66	**Sens/epoch interictal:** 0.94 ± 0.17 *Prediction:* **Sens/sz:** 0.58 **Sens/epoch:** 0.16 ± 0.18 **DL(s):** - 3.82 ± 1.85 **FPR/h:** 3.5 ± 4.02 *Detection:* **Sens/sz:** 1.0 **Sens/epoch:** 0.85 ± 0.16 **DL(s):** 3.16 ± 4.07 **FPR/h:** 2.84 ± 2.0	**Sens/epoch interictal:** 0.98 ± 0.02 *Prediction:* **Sens/sz:** 0.74 **Sens/epoch:** 0.25 ± 0.2 **DL(s):** - 6.0 ± 6.0 **FPR/h:** 7.36 ± 5.35 *Detection:* **Sens/sz:** 1.0 **Sens/epoch:** 0.5 ± 0.16 **DL(s):** 8.53 ± 7.8 **FPR/h:** 0.69 ± 1.25	2
MUN2	**Sens/epoch interictal:** 0.64 ± 0.27 *Prediction:* **Sens/sz:** 0.73 **Sens/epoch:** 0.38 ± 0.31 **DL(s):** - 6.0 ± 4.5 **FPR/h:** 61.84 ± 31.18 *Detection:* **Sens/sz:** 0.36 **Sens/epoch:** 0.19 ± 0.27 **DL(s):** 5.25 ± 1.3 **FPR/h:** 17.37 ± 29.6	**Sens/epoch interictal:** 0.77 ± 0.19 *Prediction:* **Sens/sz:** 0.64 **Sens/epoch:** 0.35 ± 0.3 **DL(s):** - 6.43 ± 4.37 **FPR/h:** 49.15 ± 33.36 *Detection:* **Sens/sz:** 0.64 **Sens/epoch:** 0.44 ± 0.35 **DL(s):** 3.43 ± 1.92 **FPR/h:** 4.94 ± 4.32	**Sens/epoch interictal:** 0.55 ± 0.26 *Prediction:* **Sens/sz:** 0.91 **Sens/epoch:** 0.53 ± 0.28 **DL(s):** - 5.4 ± 3.23 **FPR/h:** 71.63 ± 27.53 *Detection:* **Sens/sz:** 0.18 **Sens/epoch:** 0.08 ± 0.16 **DL(s):** 6.0 ± 0.0 **FPR/h:** 0.31 ± 0.65	1
MUN3	**Sens/epoch interictal:** 0.74 ± 0.28 *Prediction:* **Sens/sz:** 0.67 **Sens/epoch:** 0.37 ± 0.36 **DL(s):** - 12.75 ± 10.26 **FPR/h:** 18.71 ± 18.37 *Detection:* **Sens/sz:** 0.83 **Sens/epoch:** 0.66 ± 0.38 **DL(s):** 4.8 ± 9.6 **FPR/h:** 16.32 ± 11.71	**Sens/epoch interictal:** 0.93 ± 0.05 *Prediction:* **Sens/sz:** 0.67 **Sens/epoch:** 0.45 ± 0.38 **DL(s):** - 8.25 ± 7.46 **FPR/h:** 9.22 ± 7.62 *Detection:* **Sens/sz:** 1.0 **Sens/epoch:** 0.67 ± 0.21 **DL(s):** 5.0 ± 6.16 **FPR/h:** 12.01 ± 9.09	**Sens/epoch interictal:** 0.77 ± 0.3 *Prediction:* **Sens/sz:** 0.67 **Sens/epoch:** 0.5 ± 0.4 **DL(s):** - 8.25 ± 7.46 **FPR/h:** 24.38 ± 20.88 *Detection:* **Sens/sz:** 0.83 **Sens/epoch:** 0.56 ± 0.33 **DL(s):** 5.4 ± 9.37 **FPR/h:** 3.72 ± 3.03	2

Table 7.3: Summary of the results for EEG, ECG and multimodal fusion (LSTM), assuming 45 seconds of preictal time.

Patient	EEG	ECG	MM	Alarm size
MUN1	**Sens/epoch interictal:** 0.92 ± 0.07 *Prediction:* **Sens/sz:** 0.53 **Sens/epoch:** 0.21 ± 0.23 **DL(s):** - 12.9 ± 9.49 **FPR/h:** 9.77 ± 7.61 *Detection:* **Sens/sz:** 1.0 **Sens/epoch:** 0.61 ± 0.2 **DL(s):** 9.47 ± 9.09 **FPR/h:** 12.31 ± 13.22	**Sens/epoch interictal:** 0.97 ± 0.03 *Prediction:* **Sens/sz:** 0.47 **Sens/epoch:** 0.2 ± 0.26 **DL(s):** - 4.67 ± 2.87 **FPR/h:** 6.87 ± 7.76 *Detection:* **Sens/sz:** 1.0 **Sens/epoch:** 0.87 ± 0.15 **DL(s):** 2.37 ± 4.73 **FPR/h:** 2.29 ± 2.18	**Sens/epoch interictal:** 0.95 ± 0.04 *Prediction:* **Sens/sz:** 0.68 **Sens/epoch:** 0.34 ± 0.3 **DL(s):** - 6.69 ± 5.54 **FPR/h:** 15.31 ± 10.86 *Detection:* **Sens/sz:** 1.0 **Sens/epoch:** 0.5 ± 0.17 **DL(s):** 11.21 ± 8.97 **FPR/h:** 0.6 ± 0.7	2
MUN2	**Sens/epoch interictal:** 0.37 ± 0.23 *Prediction:* **Sens/sz:** 0.91 **Sens/epoch:** 0.64 ± 0.32 **DL(s):** - 7.5 ± 11.56 **FPR/h:** 80.03 ± 22.4 *Detection:* **Sens/sz:** 0.64 **Sens/epoch:** 0.29 ± 0.26 **DL(s):** 4.71 ± 2.19 **FPR/h:** 19.79 ± 14.53	**Sens/epoch interictal:** 0.38 ± 0.24 *Prediction:* **Sens/sz:** 0.91 **Sens/epoch:** 0.68 ± 0.26 **DL(s):** - 4.5 ± 3.61 **FPR/h:** 92.26 ± 42.89 *Detection:* **Sens/sz:** 0.91 **Sens/epoch:** 0.61 ± 0.3 **DL(s):** 2.1 ± 1.92 **FPR/h:** 33.74 ± 48.33	**Sens/epoch interictal:** 0.21 ± 0.15 *Prediction:* **Sens/sz:** 1.0 **Sens/epoch:** 0.88 ± 0.18 **DL(s):** - 3.27 ± 0.86 **FPR/h:** 57.79 ± 28.75 *Detection:* **Sens/sz:** 0.36 **Sens/epoch:** 0.16 ± 0.23 **DL(s):** 6.0 ± 2.12 **FPR/h:** 0.75 ± 1.5	1
MUN3	**Sens/epoch interictal:** 0.83 ± 0.17 *Prediction:* **Sens/sz:** 0.83 **Sens/epoch:** 0.58 ± 0.31 **DL(s):** - 9.0 ± 9.3 **FPR/h:** 30.18 ± 32.63 *Detection:* **Sens/sz:** 0.83 **Sens/epoch:** 0.57 ± 0.31 **DL(s):** 7.8 ± 8.82 **FPR/h:** 10.49 ± 7.94	**Sens/epoch interictal:** 0.88 ± 0.1 *Prediction:* **Sens/sz:** 0.83 **Sens/epoch:** 0.51 ± 0.32 **DL(s):** - 10.2 ± 7.73 **FPR/h:** 18.53 ± 14.2 *Detection:* **Sens/sz:** 1.0 **Sens/epoch:** 0.86 ± 0.13 **DL(s):** 5.0 ± 8.6 **FPR/h:** 13.5 ± 12.45	**Sens/epoch interictal:** 0.81 ± 0.16 *Prediction:* **Sens/sz:** 1.0 **Sens/epoch:** 0.68 ± 0.28 **DL(s):** - 9.0 ± 7.55 **FPR/h:** 43.7 ± 28.87 *Detection:* **Sens/sz:** 0.67 **Sens/epoch:** 0.46 ± 0.34 **DL(s):** 8.25 ± 9.34 **FPR/h:** 2.24 ± 3.2	2

For MUN2, although a high amount of seizures are predicted in the multimodal approach (91%), the false positive rate is extremely high (over 50 FP/h) for all cases, with a reduced sensitivity per seizure beyond what is clinically necessary. For this specific patient, having three classes for seizure state classification seems to be worse than having only two.

For MUN3, the best achieved sensitivity for seizure prediction is 67%, for an assumed pre-ictal period of 45s. However, the false positive rate for pre-ictal is 43.7 ± 28.87, which is not suitable for a clinical system. For 15 seconds of pre-ictal period, 50% of the seizures can be predicted, with a FPR of 4.32 ± 2.74. The average prediction time is 5.0 ± 2.8 seconds, and it is shown to increase with an increase in the assumed preictal time. Sensitivity of seizure detection is 83% with 2.36 ± 1.32 FP/h, which is a good performance.

For most cases, sensitivity of seizure prediction increases due to the multimodal approach. The false prediction rate also increases, due to the logical combination being an OR operator. However, for MUN1 and MUN3 this increase is not significant. Regarding seizure detection, the logical AND combination leads to a decrease in the false positive rate, with no significant decreases in sensitivity per seizure.

This algorithm for seizure prediction penalizes false positives, as this is modeled as a "pre-ictal state classification problems". However, false positives may be states of high seizure susceptibility where no seizures occur, and as such, may be capturing important epilepsy-related activity.

7.2.1 Discussion

These are the first results presented for multimodal seizure prediction, to the best of the author's knowledge. They validate the previously presented idea in the literature that the pre-ictal period is patient dependent, as the presented metrics vary a lot depending on the assumed length.

Here, the goal was to do an early seizure detection, instead of a long time seizure prediction, to integrate with the wearable device. This early detection should let the epileptic patient about to have a seizure move to a safer place after the warning.

Although for MUN2 the false alarm rate was too high to apply this clinically, MUN1 and MUN3 both demonstrated the feasibility of this algorithm.

Most papers in the state of the art use unimodal approaches based on intracranial EEG for seizure prediction, with higher fixed pre ictal times (varying from 10 to 40 minutes), and ECG, with pre ictal times going up to 10 minutes. As such, a higher prediction time is typically reported in those. However, some works only calculate the sensitivity for seizure prediction somewhere within the pre-ictal range, not reporting the prediction time.

This algorithm could be improved by having a more complex architecture, which gives more importance to the epochs following a pre-ictal classification, as they have a higher probability of being ictal.

Chapter 8

Conclusion and Future Work

In this dissertation, an algorithm for multimodal seizure detection and prediction was developed.

The current project is ambitious and innovative in a way that combines a multimodal integrative approach for seizure detection that can be used in different contexts (clinical and ambulatory). To the best of our knowledge, this is the first work combining more than one signal for seizure prediction, and using deep learning for multimodal seizure detection and prediction. The first steps were taken for increasing the technology readiness level of this product which may help the group acquire more funding for further developing this.

For seizure detection, after initial tests with feature extraction and machine learning (not presented) led to a lower specificity than expected using features from the state of the art, a deep learning-based approach was picked instead, for building a multimodal algorithm.

Using the presented approach, namely the pipeline including the LSTM classifier, an average of the metrics for all patients is presented in table 8.1.

Table 8.1: Mean of the results for all patients, for seizure detection, using the LSTM architecture.

EEG	**ECG**	**MM**
Sens/sz: 0.98	**Sens/sz:** 0.79	**Sens/sz:** 0.86
Sens/epoch: 0.89 ± 0.22	**Sens/epoch:** 0.77 ± 0.28	**Sens/epoch:** 0.69 ± 0.28
Spec/epoch: 0.93 ± 0.13	**Spec/epoch:** 0.89 ± 0.16	**Spec/epoch:** 0.98 ± 0.10
DL(s): 2.00 ± 6.15	DL(s): 2.14 ± 3.1	**DL(s):** 2.5 ± 3.37
FPR/h: 14.1 ± 26.3	**FPR/h:** 32.1 ± 55.1	**FPR/h:** 2.6 ± 2.6
Norm FPR/h: 44.8 ± 26.3	**Norm FPR/h:** 61.8 ± 55.15	**Norm FPR/h:** 7.9 ± 14.9

Although no direct comparison with the state of the art can be performed, as there is no patient-specific classification yet on the TUH dataset, and the Munich dataset is private, the results presented can be considered good, and show an advantage in using multimodal approaches. The shown trend is a decrease in sensitivity but with a lower number of false positives, and similar detection delays. Compared to the state of the art for multimodal approaches, sensitivity is better for some patients but the false alarm rate is a bit higher, although the same combination of sensors was not used.

For seizure prediction, the optimal pre-ictal time was shown to be patient-dependent, as previously stated - 30 seconds, non discriminative and 45 seconds, for MUN1, MUN2 and MUN3, respectively. For MUN2, the false alarm rate for detection and prediction was too high for clinical use. However, for MUN1 and MUN3, 74% and 50% of all seizures were predicted, with prediction times of 6-7 $\pm$ 5-6 seconds and 5 $\pm$ 2.8 seconds, respectively, and FPR/h of 7.4 $\pm$ 1.25 and 4.32 $\pm$ 2.74. The results for the ictal class were similar to previously presented in the chapter for seizure detection by binary classification.

Finally, we were able to demonstrate the results of the previous chapters in a portable embedded platform, NVIDIA Jetson TX1. Although the algorithm could not be validated with two wearables transmitting data at the same time, data transmission from pre-recorded EEG and ECG signals, followed by epoch classification, was successful.

Despite the low number of patients used, it is in line with the literature, as studies with similar goals presented in Table 4.6 used less than 10 patients. Moreover, it respects the number of patients needed for initial validation of a seizure detection device (1-10) according to previously defined standards (133).

The presented results show potential for further developing a wearable solution and may aid the BRAIN research group in seeking investment for continuing this line of research.

8.1 Future Work

This dissertations has shown possible paths that can be followed for developing a seizure detection and prediction system. With a few more patients from long term monitoring, the system can advance to the development of the wearable solution, which might be assisted by funding from the H2020 projects mentioned in the introduction.

However, many more studies can be performed for optimizing this system. The following subsections give some suggestions of future work.

8.1.1 Data

In spite the number of used patients being similar to some works on the state of the art, it is recommended that the number of patients is increased so the generalizing ability of the algorithm can be validated.

Moreover, the results should be presented in comparison with the type of seizure, as more patients are used, to better analyse the impact the algorithm can make in detecting different types of seizures.

8.1.2 Classification & Signal Processing

There are many more alternatives that can be used for improving classification results, such as applying different types of pre processing to the signal (e.g. band pass filters), trying different architectures and different epoch lengths (for different biosignals).

Moreover, data augmentation strategies, which are not well developed in deep learning applied to time series, may significantly reduce the false positive rate. For instance, as adding artificial eye blinks to EMG or random simulated muscle contraction to EMG and ECG may help reduce the number of false positives caused by these artifacts.

8.1.3 Online Learning & Concept Drift

On supervised learning problems in time series, concept drift can occur (134). In concept drift, there is a change in the relationships between input and output data in the underlying problem over time. This is common in physiological signals, since physiological changes occur over time. For instance, in Epilepsy Monitoring Units, due to drug deprivation, the physiological signals may be different from an ambulatory condition.

For testing if this occurs with the developed algorithm, a few patients which were admitted twice to the Epilepsy Monitoring Unit should be analyzed. The classifier should be trained on the seizures recorded on the first stay and tested on the second stay, to see if the performance deteriorated.

In parallel with this, in the ambulatory system's initial prototypes, an online learning algorithm should be implemented: when the performance degraded by a certain amount, in terms of sensitivity or false positive rate, the system should be re-trained but with a much lower learning rate. This way, the algorithm does not forget the information regarding the previous seizures and slowly updates the weights.

8.1.4 Transfer learning

The number of false positives may be reduced by pre-trained the network on a larger dataset (the combination of some of the available from the state of the art, for instance) and either train the whole network from the pre-trained weights or apply a transfer learning approach to create a patient-specific classifier. The pre-trained algorithm could also be trained on groups of patients with similar seizure characteristics.

However, there are no ECG-based public datasets of seizures, although in the group we have been collecting some of them together with S. João Hospital, for analyzing patterns of heart rate variability during the pre ictal and post ictal period. Another possible solution would be using a dataset for arrhythmia classification, which is not exactly the same problem but has some similarities, which could be learned by the network.

8.1.5 Wearable Solution

Although NVIDIA Jetson TX1 showed good computing capacity, it is very bulky and not wearable, although it is easily portable. After the initial proof of concept prototype, an investment in smaller devices (e.g. NVIDIA Jetson Nano, for around 85€) is suggested, for developing a smaller prototype that can be integrated with the developed wearable devices by BRAIN Lab (such as VitalSticker).

8.1.6 Addition of EMG to the multimodal solution

There are few studies done on EMG-based seizure detection, all of which are detailed on table 8.2. Current approaches are most effective in detecting generalized tonic-clonic seizures. As EMG can only detect motor seizures, it cannot be used for seizure prediction.

The false alarm rate in wearable devices using sEMG is still very high, although motor seizures can be well identified (133). It would be interesting to study the combination of EMG with EEG and/or ECG, to be able to detect seizures from patients with different seizure types.

Table 8.2: EMG-based seizure detection algorithms.

Reference	Algorithm	Population	Types of sz.	Avg performance	Innovation
(39)	High pass Butterworth filter (cut-off 150Hz); Count zero crossings above and below a hysteresis; Threshold	11pt/22sz	Tonic-clonic	Sens: 100%; DL: 13.7s FAR: 0.04/h	One of the pioneers of EMG-based seizure detection.
(133)	Approach in (39) (with patient-specific threshold)	71 pt /32sz	Tonic-clonic	Sens: 94%; DL: 9s; FAR: 0.03/h	Validation study for a wearable device
(42)	Feature extraction; Random Forest Classifier (patient-specific)	6 pt/27sz	Tonic-clonic	Sens: 100% (4/6 patients) FAR: 0.08-7.9/h	AI-based algorithm for seizure detection.
(41)	Band-pass filtering; Measure Maximum Voluntary Biceps contraction; Threshold-based algorithm	33/196sz 1,399h	GTCS (21) Myoclonic (96) Tonic (28) Absence (12) Focal (42)	Sens: 95% (GTCS) DL: 15.2s FAR: one false alarm (patient was restricted)	High variety of patients and low false alarm rate.

Abbreviations: GTCS: Generalized Tonic Clonic Seizure; sz: seizure; pt: patient.

8.1.7 European Projects

As previously mentioned in the introduction, our group has prepared two submissions for H2020 projects. The concept of the project is shown on Figure 8.1.

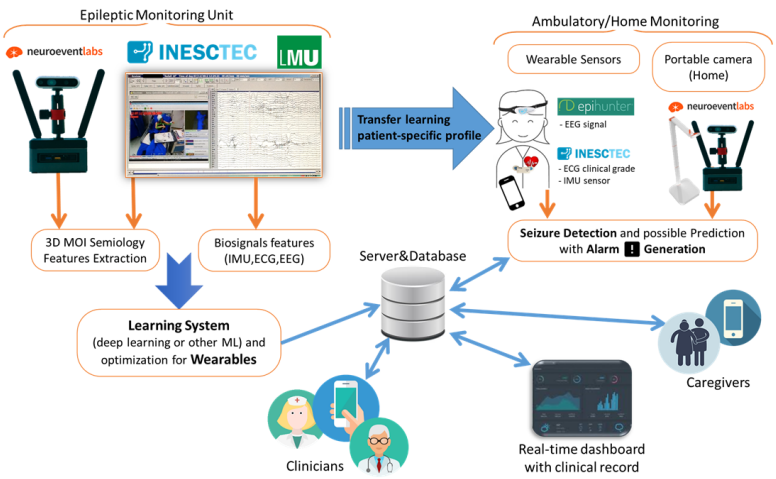


Figure 8.1: Project architecture submitted to H2020 calls (figure developed by BRAIN Group members).

The first section of the project is focused on the EMUs, where the patient is monitored and all the data (both hospital devices and wearables) is acquired under clinical observation for events and

data validation. After this, all the data is used in a learning phase where artificial intelligence is used to identify the multi-parametric physiology patterns of individual-specific epileptic seizures (e.g. acceleration/motion, EEG, ECG, 3Dvideo from Neuro Event Labs).

The second section is focused on the ambulatory/home monitoring, where an "optimized" seizure detection learned profile achieved in the EMU is used to perform near real-time data analysis using the data acquired from the patient wearable platform for precise and high-performance seizure detection. The wearables will be able to collect EEG (epihunter) and ECG waveform (INESC TEC) and movement information using an Inertial Measuring Unit (INESC TEC/LMU) or in indoor situations (including nocturnal scenarios), using the 3D camera (NEL).

If these projects are funded, this dissertation's algorithms and prototypes may be further developed to create a product with high technology readiness level (>8).

Appendix A

Features from state of the art

Table A.1: Features extracted from iEEG signals for seizure detection and prediction, in key papers chosen.

Reference	Method	Features used
		Seizure detection
(83)	SVM Classifier	Spectral energy from each channel
(110)	Random Forest Classifier	FFT Coefficients from 1 to 47Hz Correlation between channels and corresponding eigenvalues (time and frequency)
		Seizure prediction
(88)	SVM Classifier	Absolute and relative spectral power in each EEG band, spectral power ratio
(135)	Ensemble of classifiers	Spectral power, distribution statistics, AR error, fractal dimensions, correlation Hurst exponent, Riemannian autocorrelation, cross-frequency coherence,
(111)	SVM Classifier	Phase Correlation

Table A.2: Features extracted from surface EEG signals for seizure detection and prediction, in key papers chosen.

Reference	Method	Features used
		Seizure detection
(109)	Adaptive threshold	Integrated band power
		Seizure prediction
(86)	MLP Classifier	AR model predictive error, Decorrelation time, Energy, mean, variance, skewness, kurtosis Hjort complexity and mobility, relative power in each band, spectral edge frequency and power Energy of Wavelet coefficients (level 1 to 6) using Daubechies 4 mother wavelet
(66)	RNN Classifier	Relative power in each band, mean frequency, peak frequency, bandwidth of peak frequency Amplitude and duration of half waves, Spectral entropy, fractal dimension Variance, skewness, fluctuation intensity of wavelet coefficients at 2 to 5 scales of detail, and level 5 of approximation, wavelet subband entropy, autoregressive residual errors

Table A.3: Features extracted from ECG signals for seizure detection and prediction, in key papers chosen.

Reference	**Method**	**Features used**
		Seizure detection
(51)	SVM Classifier	Peak HR, HRI length, Max. HR gradient during HRI, SDNN
(114)	SVM Classifier	Peak HR, Max. HR gradient during HRI
(116)	Threshold	CSI100
(91)	SVM Classifier	SDNN, RMSSD, LF, HF, Sample Entropy, CSI, CVI
(117)	Clustering	5 non-zero eigenvalues of PCA applied on QRS Mean RR, SDNN, Average of rectified signal
		Seizure prediction
(118)	Threshold	Mean RR (ms), LF Power, HF Power, LF/HF, SD1, SD2, SD2/SD1
(82)	SVM Classifier	Mean RR, RMSSD, SDNN, NN50, PNN50, Variance(mean RR) SD1, SD2, CSI, CSV, Sample Entropy,Katz Fractal Dimension LFn, HFn, LF/HF, Recurrence Rate, Determinism, Trapping Time Length of the longest diagonal line in Recurrence Plot, Laminarity
(90)	Statistical	Mean RR, SDNN, RMSSD, Variance(RR), NN50, LF, HF, LF/HF

Appendix B

Paper accepted in MEDICON International Conference

Multimodal approach for epileptic seizure detection in Epilepsy Monitoring Units

Paulo Maia[1], *Student Member, IEEE*, Elodie Lopes[1], Elisabeth Hartl[2], Christian Vollmar[2], Soheyl Noachtar[2], and Joao Paulo Silva Cunha[1], *Senior Member, IEEE*

[1] INESC-TEC and Faculty of Engineering of University of Porto
[2] Ludwig Maximillian University of Munich

Abstract. Epilepsy is one of the most common neurological disorders, affecting up to 1% of the World population. Patients with epilepsy may suffer from severe consequences from seizures (e.g. injuries) when not monitored. Automatic seizure detection systems could mitigate this problem, improving seizure tracking and alerting a caregiver during a seizure. Existing unimodal solutions for seizure detection, based on electroencephalogram (EEG) and electrocardiogram (ECG) still have an unacceptable level of false positives, which can be reduced by combining these two biosignals. In this paper, EEG and ECG data from 7 epileptic patients with diverse recording length and seizure types were used for analyzing the importance of multimodal seizure detection, at a total of 187h46m. A leave one seizure out cross validation was selected, grouping data containing the period before a seizure and the seizure period. A proof of concept of multimodal seizure detection which uses a deep learning architecture directly on raw data is performed - a Fully Convolutional Neural Network and an architecture based on LSTM were tested. The network based on LSTM achieved better performance - using the best of one or a combination of both signals, all patients had above 91% detected seizures, a specificity per epoch above 0.96 ± 0.06 and a detection delay below 8.5 ± 12 seconds. These results show potential for developing a patient-specific approach for seizure detection that can be transferred to the ambulatory.

Keywords: epilepsy, multimodal seizure detection

1 Introduction

1.1 Background

Epilepsy is one of the most common neurological disorders, affecting from 0.5 to 1% of the world population [1]. It is characterized by a predisposition to generate epileptic seizures, a transient and temporary occurrence of signs/symptoms which are consequence of an abnormal neuronal activity in the brain [2].

Currently, for a good diagnosis and therapy planning, physicians need to admit patients into specialized centres with an Epilepsy Monitoring Unit (EMU) where continuous video-electroencephalography (vEEG) is carried out [3].

The required treatment for epilepsy is dependent on seizure type and the diagnosed epilepsy syndrome. As such, the available treatments focus on the suppression of the seizures [4], which is mainly done using anti epileptic drugs. Around 60 % of the patients become seizure-free with up to three anti-seizure medication trials [5]. However, about 25% of the patients suffering from epilepsy continue having more than one seizure per month [6], 12% and 8% have weekly and daily seizure frequency, respectively. Besides, the mortality rate of patients with active epilepsy is 5 times higher than those which are seizure free, number which has remained stable despite the introduction of new antiepileptic drugs [7, 8].

Outside EMUs, there may be severe consequences associated with seizures, such as status epilepticus (seizures with a duration over 5 minutes), injury or even death from falls or other accidents [9].

In EMUs, patient safety can be ensured by continuous observation through trained personnel. However, only about 60% of EMUs can provide this continuous patient surveillance, and around 17% of EMUs use online seizure detection and warning systems. Detection systems could optimize data review and make personnel assignment more efficient [10].

Moreover, seizure tracking, which is relevant for studying the effectiveness of an antiepileptic drug, is dependent on a subjective patient and family recall and may be influenced by the capacity of remembering details post-seizure, by the level of awareness during the seizure and by the ability to identify seizures [11]. Accuracy of seizure detection, if based on videoEEG, can be improved by 29-fold and 7-fold, respectively, if observed by families or nurses [12].

Based on these challenges, a wearable device for automatic seizure detection which can be used in both EMUs and ambulatory is of utmost interest. The best solution for seizure detection should include EEG and other physiological signals, depending on a patient's seizure type.

1.2 State of the art & Proposal

For seizure detection, depending on the actual state of the epoch being classified, one out of four outcomes are possible: a true-seizure period is missed (false negative - FN), correctly detected (true positive - TP), wrongly flagged (false positive - FP) or rejected if it is not a seizure period (true negative - TN). Three metrics are typically calculated for the first phase of validation of seizure detection algorithms: detection latency, sensitivity and false positive per hour [13].

The detection latency (DL) is the time interval from seizure onset to actual detection, and as such, is directly related to detection speed.

Sensitivity (Eq. 1) expresses the likelihood that a true seizure is detected. Sensitivity can be calculated either by seizure epoch or by seizure. When calculated by seizure, it can be done using a test set or with a grouped cross validation approach.

$$Sens = \frac{TP}{TP + FN} \tag{1}$$

The False Positive Rate (FPR) is defined as the number of false positives occurring per unit time (typically, by hour or by 24 hours). Specificity (Eq. 2) may be calculated instead, for estimating the potential of having false positives, for shorter recordings, since it is a normalized value.

$$Spec = \frac{TN}{FP + TN} \tag{2}$$

Many solutions for seizure detection are based on a single sensor (unimodal). However, unimodal seizure detectors still have an unacceptable amount of false alarms, and are not able to detect all types of seizures, as different seizures trigger different physiological responses [14]. Multimodal algorithms might help lower this false alarm rate - for instance, a non-motor seizure shows no motor changes, and as such, a seizure detector based on electromiography (EMG) or accelerometer would not work for these types of seizures. In fact, a 30-57% reduction of the false alarm rate was observed when comparing multimodal and unimodal approaches [15].

On PubMed, a search for at least one of the keywords 'seizure detection', 'seizure prediction', 'EEG seizure detection and prediction', 'EMG seizure detection', 'ECG seizure detection and prediction' returns a total of 4177 results. Of those, table 1 shows the most relevant examples of multimodal seizure detection algorithms. The major difference between multimodal methods is the way the biosignals are integrated with each other for achieving the final result, such as late fusion, which uses one classifier for each signal [16] and fuses the result using logical operations, or early fusion [17], in which the features from each signal are concatenated.

Machine learning approaches are very target-directed, and there is no way of knowing *a priori* which feature extraction algorithm is best for a given problem. A certain feature is considered to be good based on experimental approaches. As such, it requires a long time to optimize the best features for a given problem [20]. In the past few years, deep learning techniques have achieved very good results in challenges that conventional machine learning techniques could not solve. This is due to an increase in computing power and the elimination of the need for feature extraction and selection. Besides, deep learning algorithms learn the optimal features intrinsically, considering all the available evidence, leading to a more natural and unbiased representation [20].

Although deep learning methodologies are commonly used in biosignals other than EEG [20], there are no studies regarding its performance in epileptic seizure detection nor in multimodal approaches. This work shows a proof of concept on 7 patients for a personalized algorithm which uses electrocardiogram (ECG) and EEG for epileptic seizure detection. To the best of our knowledge, this the first deep learning approach which combines these two biosignals for seizure detection.

Table 1: State of the art on multimodal seizure detection algorithms.

Reference	Sensor(s)	Algorithm	Population & Data	Avg performance	Innovation
[18]	EEG EMG ECG	*EEG:* Rhythmic pattern detection compared to average spectrum; *EMG:* Line Length compared to baseline; *ECG:* CSI100 over 100 beats; Threshold; Late fusion (*OR*)	92 pt 494sz/11 978h	Sens: 86% FAR: 0.7/h	High sensitivity using reduced electrode montages.
[17]	EDA Accelerometer	*Acc:* time/frequency domain features non-linear features; *EDA:* time domain features; SVM classifier (early fusion).	80 pt (7 with sz.) 16sz/4,213h	FAR: 0.03/h Sens: 94%	Novel methodology for generalized tonic clonic sz. detection.
[19]	*First stage:* HR, SpO2 and EDA *Second stage:* HR, SpO2, EDA, ECG *Third stage:* EEG	*First stage:* Search for increases in HR and EDA and decreases in SpO2 *Second stage:* Personalized parameters for each patient *Third stage (for some subjects):* EEG-feature extraction; 3-Channel selection based on best features; kNN classifier.	10 pt 26sz/339h	*First stage:* Sens: 100% (7/10 pat., 10 sz) *Second stage:* Sens: 100% (6/10 pat., 11 sz) *Third stage:* Sens: 100% (2/3 pat)	Wearable approach concerning a patient-specific biosignal selection
[15]	Accelerometer EMG ECG	*ECG:* time domain features; *Acc.:* frequency and time domain features; *EMG:* time domain features; SVM classifier; Logical OR on both alarms, in a certain time window	7 pt 22 sz/224h	ECG+Acc. (best) Sens: 90,9% FAR: 0.08/h	Multimodal sensors lead to 75% less false alarms; ACM changes are slower than EMG;

Abbreviations: SpO2: Oxygen Saturation; sz: seizure; Acc: Accelerometer; SVM: Support Vector Machine; CSI100: Cardiac Sympathetic Index; pt: patient; sz: seizure; FAR: False Alarm Rate; Sens: Sensitivity; Spec: Specificity

2 Methods

2.1 Dataset

In a joint research line between University of Porto and the Ludwig Maximillian University of Munich, we have been working on 3D Video-EEG, and have been building the largest multimedia database of this kind. Currently, the database holds more than 300 seizures from over 100 patients [21]. Besides RGB video and 3D infrared radar, we store long term surface EEG and ECG signals synchronously recorded at 256Hz and stored. Three minutes post seizures were removed from this dataset as this is not representative of natural monitoring conditions.

The rest of the dataset was taken from Temple University Hospital [22] (TUH-EEG Corpus), is an ongoing data collection effort that has recently released a training set containing 1327 seizures from 130 patients, of which the irrelevant signal parts were manually removed by healthcare professionals.

A subset of 7 patients from both Munich and TUH-EEG was chosen to represent different seizure types, lengths and recording durations, further detailed in Table 2.

2.2 Dataset pre-processing

EEG and ECG data (256Hz) was split in three seconds segments. Each segment was labeled as ictal (seizure period) or interictal (non-seizure period) depending on the majority of samples of each class.

Table 2: Detailed information of the 7 patients used for this study.

Patient ID	Sz. amount	Duration	Interictal/ictal ratio	Syndrome	Seizure type(s)
MUN1	19	37h	182	Temporal Lobe Epilepsy	Hypermotor to generalized tonic-clonic Bilateral tonic to automotor
MUN2	11	24h04m	759	Paracentral Epilepsy	Aura to Tonic Clonic Seizure
MUN3	6	47h40m	1037	Frontal Lobe Epilepsy	Aura to Complex Motor
TUH1	8	21m41s	1.6	Unknown	Tonic-clonic seizures
TUH2	11	22m43s	19.7	Frontal Lobe Epilepsy	Absence seizures
TUH3	17	24m54s	2.17	Occipital Lobe Epilepsy	Complex partial seizures
TUH4	12	12m25s	0.58	Temporal Lobe Epilepsy	Generalized seizures

Channel selection was performed for EEG, as it can significantly reduce the computational cost of the algorithm while also decreasing the amount of irrelevant information the classifier needs to learn. In EEG recordings, if no information regarding seizure location was available, the four channels with highest variance during the ictal period were selected, as previously recommended in the literature [23]. However, if healthcare professionals had previously done the task of labeling the channels involved in seizures (or the epilepsy syndrome), manually selected channels from that region were used, to decrease the source of errors. As such, in all cases except TUH1, channels were manually selected.

2.3 Classification

Two different architectures were tested: Fully Convolutional Neural Networks (FCNs) and a network based on Long Short-Term Memory Networks (LSTM).

FCNs were recently proposed for classifying univariate time series [24], and were chosen here as a classifier due to their efficacy for time series classification in several benchmark datasets.

The used FCN architecture is built by three convolutional blocks, each with three operations: convolution, batch normalization and ReLu activation. Then, a global average pooling (GAP) is applied, averaging the results of the last convolutional block. A sigmoid function is then fully connected to GAPs output. Regarding filter size, the convolutions stride is equal to 1, with zero padding applied. The convolutions contain 128, 256 and 128 filters, respectively, with filter lengths equal to 8, 5 and 3. Figure 1 shows a representation of the FCN network used.

The LSTM architecture was previously proposed for seizure detection in an EEG dataset [25]. The data samples are input into a fully connected LSTM layer of 100 neurons, as a way of learning the short and long term dependencies between segments in each signal. The interest of such an architecture is harnessing the LSTMs' ability of learning and extracting features across the time steps in a window of 3 seconds, and not across windows. Afterwards, a time-distributed Dense layer of 50 units is applied, to calculate the cost function on all timesteps. Finally, the output of the dense layer is passed through a 1D average pooling layer, which is then input into a final sigmoid layer.

During training, a grouped leave one seizure out cross validation was performed. Epochs were grouped from the end of a seizure to the end of the follow-

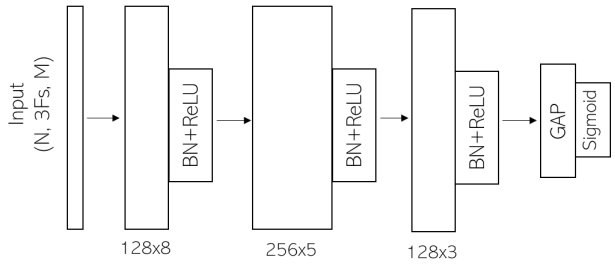


Fig. 1: Architecture of the Fully Convolutional Neural Network, extracted from [24]. In the input layer, N is the number of epochs used for training, 3Fs is the sampling frequency multiplied by three seconds and M is the number of channels (1 for ECG and 4 for EEG).

ing seizure (Fig. 2). As such, in cross validation it is guaranteed that temporally close epochs from the same seizure are not in the train and test folds, which could give an inflated estimate of model performance.

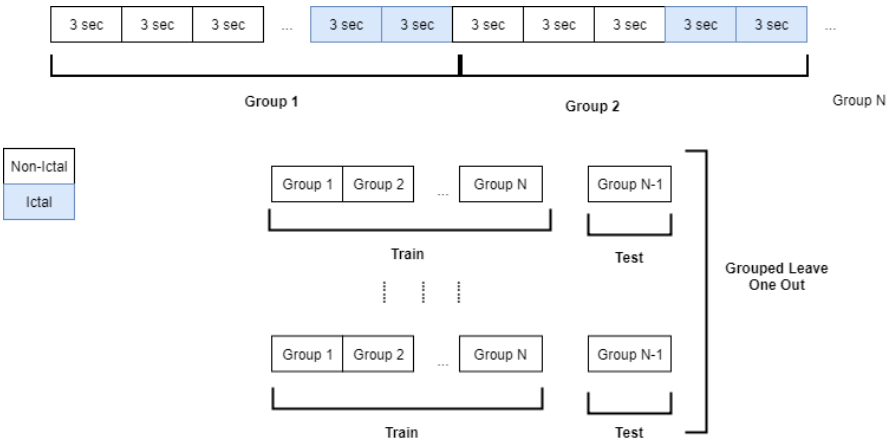


Fig. 2: Cross validation strategy for seizure detection and prediction.

In each fold of the training set, the classifier's weights were set to random and the interictal period was downsampled to be eight times the ictal period, decreasing the effect of class imbalance. In the cases where the interictal to ictal ratio was not higher than 8 (see Table 2), downsampling was performed so that the ratio between both classes was 1. These values were determined iteratively, by cross validation. In the validation set, the original size segment was used. This realistically simulates a production environment, as the model is intended to be

trained in the Epilepsy Monitoring Unit on available seizure data, and should have similar performance for detecting other seizures in a longer period, where theres a higher percentage of interictal time.

Data augmentation was applied to each training epoch as a way to increase training data - random noise with a standard deviation of 0.05 was added, and a random scaling with a normal distribution having standard deviation of 0.01 was applied. These values were selected as they did not lead to significant changes in the signal's shape (by visual analysis) and led to a better model convergence and lower false positive rate during cross-validation.

The network was trained for a maximum of 100 epochs or after the loss stopped decreasing by a factor of 0.05 for 40 epochs. Adam optimizer [26] was used with an initial learning rate of 0.001 for FCN and 0.01 for LSTM. The learning rate was decreased by a factor of 0.5 when the loss stopped decreasing for 5 epochs, up to a minimum of 0.0001. The model with the highest geometric mean between specificity and sensitivity in the validation set was then chosen as the best model. Due to class imbalance, classes were weighted accordingly to the ratio of interictal to ictal samples, and the binary crossentropy loss was used.

The output of the classifier was post-processed, and considered to be ictal if N consecutive epochs were classified as ictal, where N is patient-dependent, as different seizure types have different lengths and patterns. Depending on how much of the signal we need to see before making the optimal decision, the detection latency varies. As N is picked according to the cross-validated output, this decreases the additional bias of the classifier.

To obtain a multimodal approach, the logical AND operator was used in the output of all classifiers before post-processing.

3 Results

Table 3 and 4 summarize the obtained results for each patient, for the FCN and the LSTM architecture, respectively. As the seizure group duration is very variable, specificity was used instead of false positive rate, since it is a relative value. The sensitivity per epoch and per seizure, the detection latency and the number of consecutive epochs required to trigger a seizure alarm are also reported.

4 Discussion

This work is a proof of concept of multimodal seizure detection which uses a deep learning architecture directly on raw data, with no pre processing besides channel selection. A low number of patients was used, although it is in line with the literature, as studies with similar goals presented in Table 1 used less than 10 patients. Moreover, it respects the number of patients needed for initial validation of a seizure detection device (1-10) according to previously defined standards [13].

On average, high sensitivity values, similar to the state of the art, are obtained for detecting seizures. For classification per epoch, sensitivity is more

Table 3: Summary of the results for EEG, ECG and multimodal fusion (FCN).

Patient	EEG	ECG	MM	Alarm size
MUN1	**Sens/sz:** 1.00 **Sens/epoch:** 0.78 ± 0.15 **Spec/epoch:** 0.99 ± 0.04 **DL (s):** 3.6 ± 4.1	**Sens/sz:** 0.79 **Sens/epoch:** 0.46 ± 0.29 **Spec/epoch:** 0.99± 0.01 **DL (s):** 4.8 ± 6.1	**Sens/sz:** 0.79 **Sens/epoch:** 0.4 ± 0.28 **Spec/epoch:** 1.00 ± 0.0 **DL (s):** 7.4 ± 9.2	2
MUN2	**Sens/sz:** 0 **Sens/epoch:** 1.00 ± 0 **Spec/epoch:** 1.00 ± 0 **DL (s):** -	**Sens/sz:** 0.63 **Sens/epoch:**0.37 ± 0.32 **Spec/epoch:** 0.98 ± 0.04 **DL (s):** 3.9 ± 3.09	**Sens/sz:** 0 **Sens/epoch:** 0 **Spec/epoch:** 1 ± 0 **DL (s):** -	1
MUN3	**Sens/sz:** 0.33 **Sens/epoch:** 0.174 ± 0.33 **Spec/epoch:** 1.00 ± 0.001 **DL (s):** 7.5 ± 1.5	**Sens/sz:** 0.67 **Sens/epoch:** 0.43± 0.36 **Spec/epoch:** 0.96 ± 0.03 **DL (s):** 1.5 ± 1.5	**Sens/sz:** 0.17 **Sens/epoch:** 0.05 ± 0.11 **Spec/epoch:** 1 ± 0 **DL (s):** 12 ± 0	1
TUH1	**Sens/sz:** 1.00 **Sens/epoch:** 0.95 ± 0.03 **Spec/epoch:** 0.98 ± 0.05 **DL (s):** 1.5 ± 1.5	**Sens/sz:** 1.00 **Sens/epoch:** 0.94 ± 0.03 **Spec/epoch:** 0.99 ± 0.03 **DL (s):** 4.25 ± 2.2	**Sens/sz:** 1.00 **Sens/epoch:** 0.91 ± 0.03 **Spec/epoch:** 1.0 ± 0.00 **DL (s):** 4.25 ± 2.2	2
TUH2	**Sens/sz:** 0.9 **Sens/epoch:** 0.9 ± 0.3 **Spec/epoch:** 0.84 ± 0.29 **DL (s):** 0	**Sens/sz:** 0.9 **Sens/epoch:** 0.77 ± 0.33 **Spec/epoch:** 0.89 ± 0.29 **DL (s):** 0.33 ± 0.94	**Sens/sz:** 0.91 **Sens/epoch:** 0.77 ± 0.33 **Spec/epoch:** 0.9 ± 0.3 **DL (s):** 0.33 ± 0.94	1
TUH3	**Sens/sz:** 0.94 **Sens/epoch:** 0.45 ± 0.23 **Spec/epoch:** 0.94 ± 0.11 **DL (s):** 2.65 ± 2.34	**Sens/sz:** 0.71 **Sens/epoch:** 0.48 ± 0.41 **Spec/epoch:** 0.65 ± 0.42 **DL (s):** 6.5 ± 12.84	**Sens/sz:** 0.65 **Sens/epoch:** 0.26 ± 0.24 **Spec/epoch:** 0.96 ± 0.11 **DL (s):** 3.6 ± 2.81	1
TUH4	**Sens/sz:** 1.00 **Sens/epoch:** 0.58 ± 0.22 **Spec/epoch:** 0.96 ± 0.23 **DL (s):** 4.2 ± 1.5	**Sens/sz:** 1.00 **Sens/epoch:** 0.60 ± 0.3 **Spec/epoch:** 1 **DL (s):** 3 ± 2.3	**Sens/sz:** 1.00 **Sens/epoch:** 0.44 ± 0.27 **Spec/epoch:** 1 **DL (s):** 4.2 ± 1.5	2

Table 4: Summary of the results for EEG, ECG and multimodal fusion (LSTM).

Patient	EEG	ECG	MM	Alarm size
MUN1	**Sens/sz:** 1.00 **Sens/epoch:** 0.85 ± 0.12 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 2.5 ± 3.8	**Sens/sz:** 1.00 **Sens/epoch:** 0.55 ± 0.2 **Spec/epoch:** 0.96 ± 0.05 **DL (s):** 5.55 ± 6.1	**Sens/sz:** 1.00 **Sens/epoch:**0.5 ± 0.19 **Spec/epoch:** 0.99 ± 0.01 **DL (s):**6.3 ± 5.9	2
MUN2	**Sens/sz:** 1.00 **Sens/epoch:** 0.79 ± 0.18 **Spec/epoch:** 0.96 ± 0.01 **DL (s):** 1.6 ± 1.5	**Sens/sz:** 1.00 **Sens/epoch:**0.65 ± 0.19 **Spec/epoch:** 0.92 ± 0.08 **DL (s):** 3 ± 2.56	**Sens/sz:** 0.91 **Sens/epoch:** 0.55 ± 0.27 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 3.6 ± 2.3	1
MUN3	**Sens/sz:** 1.00 **Sens/epoch:** 0.57 ± 0.29 **Spec/epoch:** 0.97 ± 0.02 **DL (s):** 9 ± 13	**Sens/sz:** 0.83 **Sens/epoch:** 0.6 ± 0.32 **Spec/epoch:** 0.97 ± 0.02 **DL (s):** 2.4 ± 2.25	**Sens/sz:** 0.83 **Sens/epoch:** 0.44 ± 0.26 **Spec/epoch:** 0.99 ± 0.01 **DL (s):** 4.8 ± 5.6	2
TUH1	**Sens/sz:** 1.00 **Sens/epoch:** 0.93 ± 0.08 **Spec/epoch:** 0.97 ± 0.07 **DL (s):** 3 ± 1.7	**Sens/sz:** 1.00 **Sens/epoch:** 0.91 ± 0.08 **Spec/epoch:** 1 **DL (s):** 5 ± 4.1	**Sens/sz:** 1.00 **Sens/epoch:** 0.88 ± 0.12 **Spec/epoch:** 1.0 ± 0.00 **DL (s):** 5.8 ± 3.9	2
TUH2	**Sens/sz:** 0.91 **Sens/epoch:** 0.88 ± 0.29 **Spec/epoch:** 0.87 ± 0.28 **DL (s):** 0.3 ± 0.9	**Sens/sz:** 0.91 **Sens/epoch:** 0.83 ± 0.31 **Spec/epoch:** 0.96 ± 0.06 **DL (s):** 0.3 ± 0.9	**Sens/sz:** 0.91 **Sens/epoch:** 0.83 ± 0.31 **Spec/epoch:** 0.91 ± 0.29 **DL (s):** 0.3 ± 0.9	1
TUH3	**Sens/sz:** 1.00 **Sens/epoch:** 0.5 ± 0.23 **Spec/epoch:** 0.96 ± 0.06 **DL (s):** 8.5 ± 12	**Sens/sz:** 0.76 **Sens/epoch:** 0.3 ± 0.25 **Spec/epoch:** 0.93 ± 0.12 **DL (s):** 9.2 ± 5.3	**Sens/sz:** 0.65 **Sens/epoch:** 0.2 ± 0.2 **Spec/epoch:** 0.9 ± 0.02 **DL (s):** 9.5 ± 4.09	2
TUH4	**Sens/sz:** 1.00 **Sens/epoch:** 0.73 ± 0.26 **Spec/epoch:** 0.92 ± 0.15 **DL (s):** 1.2 ± 2	**Sens/sz:** 1.00 **Sens/epoch:** 0.68 ± 0.28 **Spec/epoch:** 0.79 ± 0.3 **DL (s):** 2.7 ± 4.5	**Sens/sz:** 1.00 **Sens/epoch:** 0.46 ± 0.24 **Spec/epoch:** 0.99 ± 0.04 **DL (s):** 3.6 ± 4.4	1

variable. However, what is most important is detecting a seizure and triggering an alarm, which may prevent life-threatening conditions, and not estimating its full duration. Specificity shows an increase in multimodal approaches, on average, but with a decrease in sensitivity, and a higher variability in sensitivity per epoch, in some cases. It is preferable to have high sensitivity in difficult cases, with motor symptoms which may lead to injuries, such as tonic clonic seizures.

The architecture based on LSTM shows more potential for detecting seizures, as all patients have sensitivities per seizure above 80%, while for FCN the minimum value is 33%. As such, LSTM was chosen to be the best classifier for this subset.

For patient TUH1, both EEG and ECG approaches showed relevant results as unimodal detectors. Multimodal detection did not lead to significant improvements. These results are to be expected since the patient has violent tonic-clonic seizures, easily detected using either EEG or ECG. This patient could benefit from a single wearable, such as a non intrusive ECG device, which has very high sensitivity and low detection delay, showing potential for injury prevention. Patient TUH2 shows the best sensitivity with EEG analysis, although ECG has similar results with reduced false positives, and is a less intrusive sensor and with less noise on ambulatory. Patient TUH3 has very high sensitivity per seizure in ECG, and an increased specificity in the multimodal approach, although with this data fusion methodology the sensitivity slightly decreases. As such, this patient would benefit from having a wearable with few EEG electrodes. TUH4 benefits from having a multimodal approach in terms of decrease of false positives, sacrificing sensitivity per epoch.

MUN1 shows similar behavior to TUH4. In MUN2, there is a decrease in the sensitivity per seizure in the multimodal approach, with EEG being the optimal sensor. Finally, in MUN3 the multimodal approach shows higher specificity but lower sensitivity, and also a lower delay when compared to EEG, the best biosignal for this patient.

For both MUN1 and TUH4, a high sensitivity for seizure detection using ECG is likely to happen, as temporal lobe seizures, as in 82% of the patients with epilepsy, seizures are accompanied by ictal sinus tachycardia [27], due to the activation of the central autonomic network caused by epileptic discharges. This occurs especially in patients with generalized tonic-clonic seizures and focal impaired awareness seizures, originating from the temporal lobe [28]. However, it is not clear if the seizures were detected due to tachycardia or to other factors, since this deep learning based approach has poor feature interpretability.

Data from University of Munich is more representative of realistic conditions, regarding the fact that it is composed of long term recordings, in which patients do their daily routine while in bed. On the contrary, TUH data has removed irrelevant signal segments that could be important for validating an alarm system, which is why the interictal to ictal ratio is lower than data from Munich.

In seizure detectors, the false alarm rate is dependent on the number of seizures the patient has per day. With the dataset we are developing, containing

long term recordings, we can calculate newer metrics, such as the sensitivity per hour normalized by the average distance per seizures.

In the future, more patients with different conditions and sensors (such as intracranial EEG) will be used. Moreover, complex alternatives for fusing the results from different unimodal sensors will be studied, which may help increase the performance of the classifier. These results show that logically combining the output of two signals with an AND operator is most beneficial when both signals give good results. If one seizure is not detected with one of the classifiers, the multimodal approachs sensitivity also decreases. As such, for some cases with good specificity per each biosignal, an approach based on the OR operator could be beneficial, for instance.

References

[1] Epilepsy Foundation, *https://www.epilepsy.com/learn/*, 2018.

[2] R. S. Fisher *et al.*, "Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," *Epilepsia*, vol. 46, no. 4, pp. 470–472, 2005.

[3] F. Rosenow and H. Luders, "Presurgical evaluation of epilepsy," *Brain*, 2001.

[4] E. Carrette, K. Vonck, and P. Boon, "The management of pharmacologically refractory epilepsy," *International Journal of Clinical Reviews*, 2011.

[5] P. Kwan and M. J. Brodie, "Early Identification of Refractory Epilepsy," *New England Journal of Medicine*, 2000.

[6] L. Forsgren *et al.*, *The epidemiology of epilepsy in Europe - A systematic review*, 2005.

[7] A. Neligan *et al.*, "Temporal trends in the mortality of people with epilepsy: a review," *Epilepsia*, vol. 51, no. 11, pp. 2241–2246, 2010.

[8] L. Nashef *et al.*, "Incidence of sudden unexpected death in an adult outpatient cohort with epilepsy at a tertiary referral centre," *Journal of Neurology, Neurosurgery and Psychiatry*, 1995.

[9] V. d. V. A. *et al.*, "Non-EEG seizure detection systems and potential SUDEP prevention: State of the art: Review and update," *Seizure*, 2016.

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

[11] A. Ulate-Campos *et al.*, "Automated seizure detection systems and their effectiveness for each type of seizure," *Seizure*, 2016.

[12] T. M. Nijsen *et al.*, "The potential value of three-dimensional accelerometry for detection of motor seizures in severe epilepsy," *Epilepsy and Behavior*, 2005.

[13] S. Beniczky and P. Ryvlin, "Standards for testing and clinical validation of seizure detection devices," *Epilepsia*, 2018.

[14] A. van Westrhenen *et al.*, "Ictal autonomic changes as a tool for seizure detection: a systematic review," *Clinical Autonomic Research*, pp. 1–21, 2018.

[15] T. De Cooman *et al.*, "Comparison and combination of electrocardiogram, electromyogram and accelerometry for tonic-clonic seizure detection in children," in *2018 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2018*, 2018.

[16] M. Qaraqe *et al.*, "Epileptic seizure onset detection based on EEG and ECG data fusion," *Epilepsy and Behavior*, 2016.

[17] M. Z. Poh *et al.*, "Convulsive seizure detection using a wrist-worn electrodermal activity and accelerometry biosensor," *Epilepsia*, 2012.

[18] F. Fürbass *et al.*, "Automatic multimodal detection for long-term seizure documentation in epilepsy," *Clinical Neurophysiology*, 2017.

[19] D. Cogan *et al.*, "Multi-Biosignal Analysis for Epileptic Seizure Monitoring," *International Journal of Neural Systems*, 2017.

[20] O. Faust *et al.*, "Deep learning for healthcare applications based on physiological signals: A review," *Computer Methods and Programs in Biomedicine*, 2018.

[21] H. M. Pereira Choupina *et al.*, "NeuroKinect 3.0: Multi-Bed 3Dvideo-EEG System for Epilepsy Clinical Motion Monitoring.," *Studies in health technology and informatics*, vol. 247, pp. 46–50, 2018.

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

[23] J. Duun-Henriksen *et al.*, "Channel selection for automatic seizure detection," *Clinical Neurophysiology*, 2012.

[24] Z. Wang, W. Yan, and T. Oates, "Time series classification from scratch with deep neural networks: A strong baseline," in *Proceedings of the International Joint Conference on Neural Networks*, 2017.

[25] R. Hussein *et al.*, "Epileptic Seizure Detection: A Deep Learning Approach," *arXiv preprint arXiv:1803.09848*, 2018.

[26] O. Konur, "Adam Optimizer," *Energy Education Science and Technology Part B: Social and Educational Studies*, 2013.

[27] K. S. Eggleston, B. D. Olin, and R. S. Fisher, "Ictal tachycardia: The head-heart connection," *Seizure*, 2014.

[28] C. Sevcencu and J. J. Struijk, "Autonomic alterations and cardiac changes in epilepsy," *Epilepsia*, 2010.

Appendix C

Workshop certificate MDevNet

Certificado de Participação

Certifica-se que

Paulo Maia

esteve presente no Workshop *"Prospeção e Identificação de Tecnologias com Elevado Potencial de Transferência para o Mercado"*, no âmbito do projeto MDevNet, realizado nas instalações da UPTEC.

Porto, 12 de fevereiro de 2019

Nuno Vargas

Cofinanciado por:

Appendix D

Protocol for extra data acquisition

PATIENT MOVEMENTS OF INTEREST PROTOCOL

Objective: to label common movements during seizures to have normal behavior, susceptible to false positives in many algorithms.

Physician initial position: The healthcare professional should stay at least one meter away from the bed, so the movement can be clearly seen. Please check if the healthcare professional is not being covered in the Kinect camera.

- Picking up mobile phone and talking
- Turning around in bed, to the left side and to the right side
- Sitting in bed
- Drinking water
- Coughing
- Nose scratching
- Getting up from bed
- Naturally using phone (i.e. texting, scrolling)
- Ictal testing protocol

References

[1] Epilepsy Foundation. https://www.epilepsy.com/learn/, 2018.

[2] Lars Forsgren, E. Beghi, A. Õun, and M. Sillanpää. The epidemiology of epilepsy in Europe - A systematic review, 2005.

[3] Robert S. Fisher, Walter Van Emde Boas, Warren Blume, Christian Elger, Pierre Genton, Phillip Lee, and Jerome Engel. Epileptic seizures and epilepsy: Definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). *Epilepsia*, 46(4):470–472, 2005.

[4] Robert S. Fisher, Carlos Acevedo, Alexis Arzimanoglou, Alicia Bogacz, J. Helen Cross, Christian E. Elger, Jerome Engel, Lars Forsgren, Jacqueline A. French, Mike Glynn, Dale C. Hesdorffer, B. I. Lee, Gary W. Mathern, Solomon L. Moshé, Emilio Perucca, Ingrid E. Scheffer, Torbjörn Tomson, Masako Watanabe, and Samuel Wiebe. ILAE Official Report: A practical clinical definition of epilepsy. *Epilepsia*, 2014.

[5] Ingrid E. Scheffer, Samuel Berkovic, Giuseppe Capovilla, Mary B. Connolly, Jacqueline French, Laura Guilhoto, Edouard Hirsch, Satish Jain, Gary W. Mathern, Solomon L. Moshé, Douglas R. Nordli, Emilio Perucca, Torbjörn Tomson, Samuel Wiebe, Yue Hua Zhang, and Sameer M. Zuberi. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia*, 2017.

[6] Robert S. Fisher, J. Helen Cross, Jacqueline A. French, Norimichi Higurashi, Edouard Hirsch, Floor E. Jansen, Lieven Lagae, Solomon L. Moshé, Jukka Peltola, Eliane Roulet Perez, Ingrid E. Scheffer, and Sameer M. Zuberi. Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology. *Epilepsia*, 2017.

[7] Khalid Alqadi, Ram Sankaraneni, Ursula Thome, and Prakash Kotagal. Semiology of hypermotor (hyperkinetic) seizures. *Epilepsy and Behavior*, 2016.

[8] F. Rosenow and H. Luders. Presurgical evaluation of epilepsy. *Brain*, 2001.

[9] S. J M Smith. EEG in the diagnosis, classification, and management of patients with epilepsy. *Neurology in Practice*, 2005.

[10] Evelien Carrette, Kristl Vonck, and Paul Boon. The management of pharmacologically refractory epilepsy. *International Journal of Clinical Reviews*, 2011.

[11] Kostas N. Fountas and J. R. Smith. A novel closed-loop stimulation system in the control of focal, medically refractory epilepsy. *Acta Neurochirurgica, Supplementum*, 2007.

[12] Jerry J. Shih, Nathan B. Fountain, Susan T. Herman, Anto Bagic, Fred Lado, Susan Arnold, Mary L. Zupanc, Ellen Riker, and David M. Labiner. Indications and methodology for video-electroencephalographic studies in the epilepsy monitoring unit, 2018.

[13] P Kwan and M J Brodie. Early identification of refractory epilepsy. *The New England journal of medicine*, 2000.

[14] Benno Mahler, Sofia Carlsson, Tomas Andersson, and Torbjörn Tomson. Risk for injuries and accidents in epilepsy: a prospective population-based cohort study. *Neurology*, 90(9):e779–e789, 2018.

[15] Barbara Blachut, Christian Hoppe, Rainer Surges, Jutta Stahl, Christian E. Elger, and Christoph Helmstaedter. Counting seizures: The primary outcome measure in epileptology from the patients' perspective. *Seizure*, 2015.

[16] Christian E. Elger and Christian Hoppe. Diagnostic challenges in epilepsy: seizure underreporting and seizure detection. *The Lancet Neurology*, 2018.

[17] A. Ulate-Campos, F. Coughlin, M. Gaínza-Lein, I. Sánchez Fernández, P. L. Pearl, and T. Loddenkemper. Automated seizure detection systems and their effectiveness for each type of seizure. *Seizure*, 2016.

[18] Tamara M.E. Nijsen, Johan B.A.M. Arends, Paul A.M. Griep, and Pierre J.M. Cluitmans. The potential value of three-dimensional accelerometry for detection of motor seizures in severe epilepsy. *Epilepsy and Behavior*, 2005.

[19] Christoph Baumgartner, Johannes P Koren, and Michaela Rothmayer. Automatic Computer-Based Detection of Epileptic Seizures. *Frontiers in neurology*, 9:639, 8 2018.

[20] Van de Vel A., Cuppens K., Bonroy B., Milosevic M., Jansen K., Van Huffel S., Vanrumste B., Cras P., and Lagae L. Non-EEG seizure detection systems and potential SUDEP prevention: State of the art: Review and update. *Seizure*, 2016.

[21] Ivan Osorio, Hitten P Zaveri, Mark G Frei, and Susan Arthurs. *Epilepsy: the intersection of neurosciences, biology, mathematics, engineering, and physics*. CRC press, 2016.

[22] Guido Rubboli, Sandor Beniczky, Steven Claus, Maria Paola Canevini, Philippe Kahane, Hermann Stefan, Walter van Emde Boas, Demetrios Velis, Elise Reus, Antonio Gil-Nagel, Bernhard J. Steinhoff, Eugen Trinka, and Philippe Ryvlin. A European survey on current practices in epilepsy monitoring units and implications for patients' safety. *Epilepsy and Behavior*, 2015.

[23] Khara M. Sauro, Natalie Wiebe, Sophie Macrodimitris, Samuel Wiebe, Sara Lukmanji, and Nathalie Jetté. Quality and safety in adult epilepsy monitoring units: A systematic review and meta-analysis. *Epilepsia*, 2016.

[24] Alla Guekht, Maria Mizinova, Igor Kaimovsky, Oksana Danilenko, Elisa Bianchi, and Ettore Beghi. The direct costs of epilepsy in Russia. A prospective cost-of-illness study from a single center in Moscow. *Epilepsy & Behavior*, 64:122–126, 2016.

[25] J. Olesen, A. Gustavsson, M. Svensson, H. U. Wittchen, and B. Jönsson. The economic cost of brain disorders in Europe. *European Journal of Neurology*, 2012.

[26] D L W Davidson and S Macdonald. The costs of trauma caused by seizures: can they be reduced? *Seizure*, 11(5):344–347, 2002.

[27] Lena-Marie Kortland, Susanne Knake, Felix Rosenow, and Adam Strzelczyk. Cost of status epilepticus: a systematic review. *Seizure*, 24:17–20, 2015.

[28] A. Tetto, P. Manzoni, A. Millul, Ettore Beghi, L. Garattini, A. Tartara, and G. Avanzini. The costs of epilepsy in Italy: A prospective cost-of-illness study in referral patients with disease of different severity. *Epilepsy Research*, 2002.

[29] M. De Zélicourt, B. De Toffol, H. Vespignani, C. Laurendeau, L. Lévy-Bachelot, C. Murat, and F. Fagnani. Management of focal epilepsy in adults treated with polytherapy in France: The direct cost of drug resistance (ESPERA study). *Seizure*, 2014.

[30] Anna H. Noda, Anke Hermsen, Ralf Berkenfeld, Dieter Dennig, Günther Endrass, Jens Kaltofen, Ali Safavi, Stephan Wiehler, Gunther Carl, Uwe Meier, Christian E. Elger, Katja Menzler, Susanne Knake, Felix Rosenow, and Adam Strzelczyk. Evaluation of costs of epilepsy using an electronic practice management software in Germany. *Seizure*, 2015.

[31] D. L. Schomer and F. Lopes da Silva. *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. 2010.

[32] George H Klem, Hans Otto Luëders, H Jasper, and C Elger. The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, 1999.

[33] Daniel Drane. Cognitive Effects of Chronic Epilepsy. *Epilepsy*, pages 260–267, 2014.

[34] Prasanna Jayakar, Jean Gotman, A. Simon Harvey, André Palmini, Laura Tassi, Donald Schomer, Francois Dubeau, Fabrice Bartolomei, Alice Yu, Pavel Kršek, Demetrios Velis, and Philippe Kahane. Diagnostic utility of invasive EEG for epilepsy surgery: Indications, modalities, and techniques. *Epilepsia*, 2016.

[35] Soheyl Noachtar and Jan Rémi. The role of EEG in epilepsy: A critical review. *Epilepsy and Behavior*, 2009.

[36] Joseph D. Bronzino. *The Biomedical Engineering Handbook*. 2000.

[37] Aashit K Shah and Sandeep Mittal. Invasive electroencephalography monitoring: Indications and presurgical planning. *Annals of Indian Academy of Neurology*, 17(Suppl 1):S89, 2014.

[38] Dean R. Freestone, Philippa J. Karoly, and Mark J. Cook. A forward-looking review of seizure prediction, 2017.

[39] Isa Conradsen, Sndor Beniczky, Karsten Hoppe, Peter Wolf, and Helge B.D. Sorensen. Automated algorithm for generalized tonic-clonic epileptic seizure onset detection based on sEMG zero-crossing rate. *IEEE Transactions on Biomedical Engineering*, 2012.

[40] Jonathan J. Halford, Michael R. Sperling, Dileep R. Nair, Dennis J. Dlugos, William O. Tatum, Jay Harvey, Jacqueline A. French, John R. Pollard, Edward Faught, Katherine H. Noe, Thomas R. Henry, Gina M. Jetter, Octavian V. Lie, Lola C. Morgan, Michael R. Girouard, Damon P. Cardenas, Luke E. Whitmire, and Jose E. Cavazos. Detection of generalized tonic–clonic seizures using surface electromyographic monitoring. *Epilepsia*, 2017.

[41] Charles Ákos Szabõ, Lola C. Morgan, Kameel M. Karkar, Linda D. Leary, Octavian V. Lie, Michael Girouard, and José E. Cavazos. Electromyography-based seizure detector: Preliminary results comparing a generalized tonic-clonic seizure detection algorithm to video-EEG recordings. In *Epilepsia*, 2015.

[42] Sigge N. Larsen, Isa Conradsen, Sandor Beniczky, and Helge B.D. Sorensen. Detection of tonic epileptic seizures based on surface electromyography. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, 2014.

[43] Isa Conradsen, Peter Wolf, Thomas Sams, Helge B.D. Sorensen, and Sándor Beniczky. Patterns of muscle activation during generalized tonic and tonic-clonic epileptic seizures. *Epilepsia*, 2011.

[44] Thomas De Cooman, Troels W Kjær, Sabine Van Huffel, and Helge B Sorensen. Adaptive heart rate-based epileptic seizure detection using real-time user feedback. *Physiological measurement*, 39(1):14005, 2018.

[45] Katherine S. Eggleston, Bryan D. Olin, and Robert S. Fisher. Ictal tachycardia: The head-heart connection. *Seizure*, 2014.

[46] Cristian Sevcencu and Johannes J. Struijk. Autonomic alterations and cardiac changes in epilepsy. *Epilepsia*, 2010.

[47] C. Baumgartner, S. Lurger, and F. Leutmezer. Autonomic symptoms during epileptic seizures. *Epileptic Disorders*, 2001.

[48] Maeike Zijlmans, Danny Flanagan, and Jean Gotman. Heart rate changes and ECG abnormalities during epileptic seizures: Prevalence and definition of an objective clinical sign. *Epilepsia*, 2002.

[49] Jesper Jeppesen, Sándor Beniczky, Anders Fuglsang-Frederiksen, Per Sidenius, and Yousef Jasemian. Detection of epileptic-seizures by means of power spectrum analysis of heart rate variability: A pilot study. *Technology and Health Care*, 2010.

[50] Elisa Bruno, Andrea Biondi, Mark P Richardson, and RADAR-CNS Consortium. Pre-ictal heart rate changes: A systematic review and meta-analysis. *Seizure*, 55:48–56, 2018.

[51] Thomas De Cooman, Carolina Varon, Borbála Hunyadi, Wim Van Paesschen, Lieven Lagae, and Sabine Van Huffel. Online automated seizure detection in temporal lobe epilepsy patients using single-lead ecg. *International journal of neural systems*, 27(07):1750022, 2017.

[52] Thomas de Cooman. *Epileptic Seizure Detection in a Home Environment*. PhD thesis, 2018.

[53] Johan B A M Arends. Movement-based seizure detection. *Epilepsia*, 2018.

[54] Sriram Ramgopal, Sigride Thome-Souza, Michele Jackson, Navah Ester Kadish, Iván Sánchez Fernández, Jacquelyn Klehm, William Bosl, Claus Reinsberger, Steven Schachter, and Tobias Loddenkemper. Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy. *Epilepsy and Behavior*, 2014.

[55] Andreas Schulze-Bonhage, Francisco Sales, Kathrin Wagner, Rute Teotonio, Astrid Carius, Annette Schelle, and Matthias Ihle. Views of patients with epilepsy on seizure prediction devices. *Epilepsy and Behavior*, 2010.

[56] Erie G. Gutierrez, Nathan E. Crone, Joon Y. Kang, Yaretson I. Carmenate, and Gregory L. Krauss. Strategies for non-EEG seizure detection and timing for alerting and interventions with tonic–clonic seizures. *Epilepsia*, 2018.

[57] Anup D. Patel, Robert Moss, Steven W. Rust, Jeremy Patterson, Robert Strouse, Satyanarayana Gedela, Jesse Haines, and Simon M. Lin. Patient-centered design criteria for wearable seizure detection devices. *Epilepsy and Behavior*, 2016.

[58] Nasser M Nasrabadi. Pattern recognition and machine learning. *Journal of electronic imaging*, 16(4):49901, 2007.

[59] Oliver Faust, Yuki Hagiwara, Tan Jen Hong, Oh Shu Lih, and U. Rajendra Acharya. Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine*, 2018.

[60] Yann LeCun and Yoshua Bengio. Convolutional Networks for Images, Speech, and Time Series. In *The Handbook of Brain Theory and Neural Networks*. 1998.

[61] Felix Gers. Long short-term memory in recurrent neural networks. *Neural Computation*, 2001.

[62] Colah. http://colah.github.io/posts/2015-08-Understanding-LSTMs/, 2018.

[63] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 1997.

[64] Ye Yuan, Guangxu Xun, Fenglong Ma, Qiuling Suo, Hongfei Xue, Kebin Jia, and Aidong Zhang. A novel channel-aware attention framework for multi-channel EEG seizure detection via multi-view deep learning. In *2018 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2018*, 2018.

[65] Pascal Vincent, Hugo Larochelle, Yoshua Bengio, and Pierre-Antoine Manzagol. Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning - ICML '08*, 2008.

[66] Georgiy R. Minasyan, John B. Chatten, Martha J. Chatten, and Richard N. Harner. Patient-specific early seizure detection from scalp electroencephalogram. *Journal of Clinical Neurophysiology*, 2010.

[67] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987.

[68] Girish Chandrashekar and Ferat Sahin. A survey on feature selection methods. *Computers and Electrical Engineering*, 2014.

[69] R P L Durgabai. Feature Selection using ReliefF Algorithm. *International Journal of Advanced Research in Computer and Communication Engineering*, 2014.

[70] Nhan Duy Truong, Levin Kuhlmann, Mohammad Reza Bonyadi, Jiawei Yang, Andrew Faulks, and Omid Kavehei. Supervised learning in automatic channel selection for epileptic seizure detection. *Expert Systems with Applications*, 2017.

[71] Christophe C. Jouny, Piotr J. Franaszczuk, and Gregory K. Bergey. Signal complexity and synchrony of epileptic seizures: Is there an identifiable preictal period? *Clinical Neurophysiology*, 2005.

[72] Ralph G. Andrzejak, Klaus Lehnertz, Florian Mormann, Christoph Rieke, Peter David, and Christian E. Elger. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 2001.

[73] Benjamin H. Brinkmann, Joost Wagenaar, Drew Abbot, Phillip Adkins, Simone C. Bosshard, Min Chen, Quang M. Tieng, Jialune He, F. J. Muñoz-Almaraz, Paloma Botella-Rocamora, Juan Pardo, Francisco Zamora-Martinez, Michael Hills, Wei Wu, Iryna Korshunova, Will Cukierski, Charles Vite, Edward E. Patterson, Brian Litt, and Gregory A. Worrell. Crowdsourcing reproducible seizure forecasting in human and canine epilepsy. *Brain*, 2016.

[74] M. Winterhalder, T. Maiwald, H. U. Voss, R. Aschenbrenner-Scheibe, J. Timmer, and A. Schulze-Bonhage. The seizure prediction characteristics: A general framework to assess and compare seizure prediction methods. *Epilepsy and Behavior*, 2003.

[75] Matthias Ihle, Hinnerk Feldwisch-Drentrup, César A. Teixeira, Adrien Witon, Björn Schelter, Jens Timmer, and Andreas Schulze-Bonhage. EPILEPSIAE - A European epilepsy database. *Computer Methods and Programs in Biomedicine*, 2012.

[76] Levin Kuhlmann, Philippa Karoly, Dean R Freestone, Benjamin H Brinkmann, Andriy Temko, Alexandre Barachant, Feng Li, Gilberto Titericz, Brian W Lang, Daniel Lavery, Kelly Roman, Derek Broadhead, Scott Dobson, Gareth Jones, Qingnan Tang, Irina Ivanenko, Oleg Panichev, Timothée Proix, Michal Náhlík, Daniel B Grunberg, Chip Reuben, Gregory Worrell, Brian Litt, David T J Liley, David B Grayden, and Mark J Cook. Epilepsyecosystem.org: crowd-sourcing reproducible seizure prediction with long-term human intracranial EEG. *Brain*, 2018.

[77] Iyad Obeid and Joseph Picone. The Temple University Hospital EEG Data Corpus. *Frontiers in Neuroscience*, 10:196, 2016.

[78] Md Kafiul Islam, Amir Rastegarnia, and Zhi Yang. Methods for artifact detection and removal from scalp EEG: A review. *Neurophysiologie Clinique/Clinical Neurophysiology*, 2016.

[79] Marwa Qaraqe, Muhammad Ismail, Erchin Serpedin, and Haneef Zulfi. Epileptic seizure onset detection based on EEG and ECG data fusion. *Epilepsy and Behavior*, 2016.

[80] Meysam Golmohammadi, Saeedeh Ziyabari, Vinit Shah, Silvia Lopez de Diego, Iyad Obeid, and Joseph Picone. Deep Architectures for Automated Seizure Detection in Scalp EEGs. *arXiv preprint arXiv:1712.09776*, 2017.

[81] M. Varanini, G. Tartarisco, L. Billeci, A. Macerata, G. Pioggia, and R. Balocchi. An efficient unsupervised fetal QRS complex detection from abdominal maternal ECG. *Physiological Measurement*, 2014.

[82] Lucia Billeci, Daniela Marino, Laura Insana, Giampaolo Vatti, and Maurizio Varanini. Patient-specific seizure prediction based on heart rate variability and recurrence quantification analysis. *PloS one*, 13(9):e0204339, 2018.

[83] Alaa Kharbouch, Ali Shoeb, John Guttag, and Sydney S. Cash. An algorithm for seizure onset detection using intracranial EEG. *Epilepsy and Behavior*, 2011.

[84] F. Fürbass, P. Ossenblok, M. Hartmann, H. Perko, A. M. Skupch, G. Lindinger, L. Elezi, E. Pataraia, A. J. Colon, C. Baumgartner, and T. Kluge. Prospective multi-center study of an automatic online seizure detection system for epilepsy monitoring units. *Clinical Neurophysiology*, 2015.

[85] Ihsan Ullah, Muhammad Hussain, and Hatim Aboalsamh. An automated system for epilepsy detection using EEG brain signals based on deep learning approach. *Expert Systems with Applications*, 107:61–71, 2018.

[86] César Alexandre Teixeira, Bruno Direito, Mojtaba Bandarabadi, Michel Le Van Quyen, Mario Valderrama, Bjoern Schelter, Andreas Schulze-Bonhage, Vincent Navarro, Francisco Sales, and António Dourado. Epileptic seizure predictors based on computational intelligence techniques: A comparative study with 278 patients. *Computer Methods and Programs in Biomedicine*, 2014.

[87] Nhan Duy Truong, Anh Duy Nguyen, Levin Kuhlmann, Mohammad Reza Bonyadi, Jiawei Yang, Samuel Ippolito, and Omid Kavehei. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Networks*, 2018.

[88] Zisheng Zhang and Keshab K Parhi. Low-complexity seizure prediction from iEEG/sEEG using spectral power and ratios of spectral power. *IEEE transactions on biomedical circuits and systems*, 10(3):693–706, 2016.

[89] Jesper Jeppesen, Sandor Beniczky, Peter Johansen, Per Sidenius, and Anders Fuglsang-Frederiksen. Using Lorenz plot and Cardiac Sympathetic Index of heart rate variability for detecting seizures for patients with epilepsy. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*, 2014.

[90] Koichi Fujiwara, Miho Miyajima, Toshitaka Yamakawa, Erika Abe, Yoko Suzuki, Yuriko Sawada, Manabu Kano, Taketoshi Maehara, Katsuya Ohta, Taeko Sasai-Sakuma, Tetsuo Sasano, Masato Matsuura, and Eisuke Matsushima. Epileptic Seizure Prediction Based on Multivariate Statistical Process Control of Heart Rate Variability Features. *IEEE Transactions on Biomedical Engineering*, 2016.

[91] Jonatas Pavei, Renan G. Heinzen, Barbora Novakova, Roger Walz, Andrey J. Serra, Markus Reuber, Athi Ponnusamy, and Jefferson L.B. Marques. Early seizure detection based on cardiac autonomic regulation dynamics. *Frontiers in Physiology*, 2017.

[92] Maria Hügle, Simon Heller, Manuel Watter, Manuel Blum, Farrokh Manzouri, Matthias Dümpelmann, Andreas Schulze-Bonhage, Peter Woias, and Joschka Boedecker. Early Seizure Detection with an Energy-Efficient Convolutional Neural Network on an Implantable Microcontroller. *arXiv preprint arXiv:1806.04549*, 2018.

[93] Thomas De Cooman, Carolina Varon, Anouk Van De Vel, Berten Ceulemans, Lieven Lagae, and Sabine Van Huffel. Comparison and combination of electrocardiogram, electromyogram and accelerometry for tonic-clonic seizure detection in children. In *2018 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2018*, 2018.

[94] Javad Birjandtalab, Maziyar Baran Pouyan, Diana Cogan, Mehrdad Nourani, and Jay Harvey. Automated seizure detection using limited-channel EEG and non-linear dimension reduction. *Computers in Biology and Medicine*, 2017.

[95] Jonas Duun-Henriksen, Troels Wesenberg Kjaer, Rasmus Elsborg Madsen, Line Sofie Remvig, Carsten Eckhart Thomsen, and Helge Bjarup Dissing Sorensen. Channel selection for automatic seizure detection. *Clinical Neurophysiology*, 2012.

[96] Y U Khan and J Gotman. Wavelet based automatic seizure detection in intracerebral electroencephalogram. *Clinical Neurophysiology*, 114(5):898–908, 2003.

[97] U Rajendra Acharya, Yuki Hagiwara, and Hojjat Adeli. Automated seizure prediction. *Epilepsy & Behavior*, 88:251–261, 2018.

[98] R. Esteller, J. Echauz, T. Tcheng, B. Litt, and B. Pless. Line length: An efficient feature for seizure onset detection. *Annual International Conference of the IEEE Engineering in Medicine and Biology-Proceedings*, 2001.

[99] Ashfaque Shafique, Mohamed Sayeed, and Konstantinos Tsakalis. Nonlinear Dynamical Systems with Chaos and Big Data: A Case Study of Epileptic Seizure Prediction and Control. In *Guide to Big Data Applications*, pages 329–369. Springer, 2018.

[100] Ardalan Aarabi and Bin He. A rule-based seizure prediction method for focal neocortical epilepsy. *Clinical Neurophysiology*, 123(6):1111–1122, 2012.

[101] Fred Shaffer and J. P. Ginsberg. An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, 2017.

[102] Lucia Billeci, Daniela Marino, Laura Insana, Giampaolo Vatti, and Maurizio Varanini. Patient-specific seizure prediction based on heart rate variability and recurrence quantification analysis. *PloS one*, 13(9):e0204339, 2018.

[103] D. Singh, K. Vinod, and S. C. Saxena. Sampling frequency of the RR interval time series for spectral analysis of heart rate variability. *Journal of Medical Engineering and Technology*, 2004.

[104] Ye Yuan, Guangxu Xun, Kebin Jia, and Aidong Zhang. A Multi-view Deep Learning Method for Epileptic Seizure Detection using Short-time Fourier Transform. In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology,and Health Informatics - ACM-BCB '17*, 2017.

[105] Yanli Zhang, Weidong Zhou, Shasha Yuan, and Qi Yuan. Seizure detection method based on fractal dimension and gradient boosting. *Epilepsy and Behavior*, 2015.

[106] Negin Moghim and David W. Corne. Predicting epileptic seizures in advance. *PLoS ONE*, 2014.

[107] Luigi Chisci, Antonio Mavino, Guido Perferi, Marco Sciandrone, Carmelo Anile, Gabriella Colicchio, and Filomena Fuggetta. Real-time epileptic seizure prediction using AR models and support vector machines. *IEEE Transactions on Biomedical Engineering*, 2010.

[108] Ramy Hussein, Hamid Palangi, Rabab Ward, and Z Jane Wang. Epileptic Seizure Detection: A Deep Learning Approach. *arXiv preprint arXiv:1803.09848*, 2018.

[109] Rüdiger Hopfengärtner, Burkhard S. Kasper, Wolfgang Graf, Stephanie Gollwitzer, Gernot Kreiselmeyer, Hermann Stefan, and Hajo Hamer. Automatic seizure detection in long-term scalp EEG using an adaptive thresholding technique: A validation study for clinical routine. *Clinical Neurophysiology*, 2014.

[110] Michael Hill. Seizure Detection Using FFT, Temporal and Spectral Correlation Coefficients, Eigenvalues and Random Forest, Github. Technical report, 2014.

[111] Mohammad Zavid Parvez and Manoranjan Paul. Epileptic seizure prediction by exploiting spatiotemporal relationship of EEG signals using phase correlation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2016.

[112] Elie Bou Assi, Dang K. Nguyen, Sandy Rihana, and Mohamad Sawan. Towards accurate prediction of epileptic seizures: A review. *Biomedical Signal Processing and Control*, 2017.

[113] Kais Gadhoumi, Jean Marc Lina, Florian Mormann, and Jean Gotman. Seizure prediction for therapeutic devices: A review, 2016.

[114] T. De Cooman, C. Varon, A. Van De Vel, B. Ceulemans, L. Lagae, and S. Van Huffel. Semi-supervised one-class transfer learning for heart rate based epileptic seizure detection. In *Computing in Cardiology*, 2017.

[115] Thomas De Cooman, Troels W Kjær, Sabine Van Huffel, and Helge B Sorensen. Adaptive heart rate-based epileptic seizure detection using real-time user feedback. *Physiological measurement*, 39(1):14005, 2018.

[116] Jesper Jeppesen, Sándor Beniczky, Peter Johansen, Per Sidenius, and Anders Fuglsang-Frederiksen. Detection of epileptic seizures with a modified heart rate variability algorithm based on Lorenz plot. *Seizure*, 2015.

[117] Carolina Varon, Katrien Jansen, Lieven Lagae, and Sabine Van Huffel. Can ECG monitoring identify seizures? In *Journal of Electrocardiology*, 2015.

[118] Soroor Behbahani, Nader Jafarnia Dabanloo, Ali Motie Nasrabadi, and Antonio Dourado. Prediction of epileptic seizures based on heart rate variability. *Technology and Health Care*, 2016.

[119] Anouk van Westrhenen, Thomas De Cooman, Richard H C Lazeron, Sabine Van Huffel, and Roland D Thijs. Ictal autonomic changes as a tool for seizure detection: a systematic review. *Clinical Autonomic Research*, pages 1–21, 2018.

[120] F. Fürbass, S. Kampusch, E. Kaniusas, J. Koren, S. Pirker, R. Hopfengärtner, H. Stefan, T. Kluge, and C. Baumgartner. Automatic multimodal detection for long-term seizure documentation in epilepsy. *Clinical Neurophysiology*, 2017.

[121] Ming Zher Poh, Tobias Loddenkemper, Claus Reinsberger, Nicholas C. Swenson, Shubhi Goyal, Mangwe C. Sabtala, Joseph R. Madsen, and Rosalind W. Picard. Convulsive seizure detection using a wrist-worn electrodermal activity and accelerometry biosensor. *Epilepsia*, 2012.

[122] Diana Cogan, Javad Birjandtalab, Mehrdad Nourani, Jay Harvey, and Venkatesh Nagaraddi. Multi-Biosignal Analysis for Epileptic Seizure Monitoring. *International Journal of Neural Systems*, 2017.

[123] Xiuhe Zhao and Samden D Lhatoo. Seizure detection: do current devices work? And when can they be useful? *Current Neurology and Neuroscience Reports*, 18(7):40, 2018.

[124] Hugo Miguel Pereira Choupina, Ana Patrícia Rocha, José Maria Fernandes, Christian Vollmar, Soheyl Noachtar, and João Paulo Silva Cunha. NeuroKinect 3.0: Multi-Bed 3Dvideo-EEG System for Epilepsy Clinical Motion Monitoring. *Studies in health technology and informatics*, 247:46–50, 2018.

[125] Zhiguang Wang, Weizhong Yan, and Tim Oates. Time series classification from scratch with deep neural networks: A strong baseline. In *Proceedings of the International Joint Conference on Neural Networks*, 2017.

[126] Kesheng Wu, Ekow Otoo, and Arie Shoshani. Optimizing connected component labeling algorithms. In *Medical Imaging 2005: Image Processing*, 2005.

[127] Ozcan Konur. Adam Optimizer. *Energy Education Science and Technology Part B: Social and Educational Studies*, 2013.

[128] Hassan Ismail Fawaz, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, and Pierre Alain Muller. Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*, 2019.

[129] Bowen Song, Guopeng Zhang, Wei Zhu, and Zhengrong Liang. ROC operating point selection for classification of imbalanced data with application to computer-aided polyp detection in CT colonography. *International Journal of Computer Assisted Radiology and Surgery*, 2014.

[130] Scott M. Lundberg, Bala Nair, Monica S. Vavilala, Mayumi Horibe, Michael J. Eisses, Trevor Adams, David E. Liston, Daniel King Wai Low, Shu Fang Newman, Jerry Kim, and Su In Lee. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature Biomedical Engineering*, 2018.

[131] Isabell Kiral-Kornek, Subhrajit Roy, Ewan Nurse, Benjamin Mashford, Philippa Karoly, Thomas Carroll, Daniel Payne, Susmita Saha, Steven Baldassano, Terence O'Brien, David Grayden, Mark Cook, Dean Freestone, and Stefan Harrer. Epileptic Seizure Prediction Using Big Data and Deep Learning: Toward a Mobile System. *EBioMedicine*, 2017.

[132] NVIDIA Jetson. http://www.nvidia.com/object/jetson-tx1-module.html., 2018.

[133] Sándor Beniczky, Isa Conradsen, Oliver Henning, Martin Fabricius, and Peter Wolf. Automated real-time detection of tonic-clonic seizures using a wearable EMG device. *Neurology*, 2018.

[134] Alexey Tsymbal. The problem of concept drift: definitions and related work. *Computer Science Department, Trinity College Dublin*, 2004.

[135] Levin Kuhlmann, Klaus Lehnertz, Mark P. Richardson, Björn Schelter, and Hitten P. Zaveri. Seizure prediction — ready for a new era. *Nature Reviews Neurology*, 2018.