UNIVERSITA' DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics of Data

Final dissertation

Deep learning models for detection and foreseeing of epileptic seizures

Thesis Supervisor:

Prof. Marco Zanetti

Internship Supervisor:

Prof. Alberto Testolin

Candidate:

Edoardo Antonaci

Matricola N. 1234431

Academic Year 2021/2022

To my family

Abstract

Epilepsy is one of the most widespread globally diseases and nowadays several steps to its classification and treatment have been made.
Starting from the electroencephalogram (EEG) of epileptic patients, the first step analysis consists of choosing those TRC files with an *effectively* recording channels into a range included between 24 and 35: the number of channels changes a lot from patient to patient, so it is *threshold* to collect as much as possible quite homogeneous data.

Converting each selected file into a dataframe, the analysis goes on with splitting them into no-ictal and ictal phases: using Brainstorm is actually possible to extract the start and the end events related to epileptic seizure (ES).
After checking that every ES is truely indipendent from the others (this happens after 4h from the end of the previous one), the time data order for each no-ictal part is reversed and every 5 seconds the signal is first downsampled to 256 Hz and then log-spectrogram is computed for each channel.

Finally a filter system composed by 1 Hz high pass, 125 Hz low pass and notch filter to 50 and 100 Hz is applied to clean the spectrogram from noise frequencies.
If some dataframe has less then 35 channels an empty spectograms with null signal are added in order to have 35 images for each 5 seconds time period.
As the aim of the work is to predict the ES in increasingly detailed time brackets, each *spectrogram-packet* (composed of 35 images) is first labeled and then the *new* spectrogram dataset is balanced between time categories.

Next, a split between the training and test set is necessary in order to train a CNN model in supervised learning. The test set is the 10 % of the entire dataset and it's composed by chosen random item from the various categories: that why it's called random test.
To better tuning the model hyperparameters a 3-fold cross validation is performed and to boost up the accuracy both training and random test are normalized with mean and std computed on training part.

Despite on random test an 88 % of accuracy is achieved, on the last two patients benchtest the performance falls down near to 28 %.
So, using a dataset with unknown information about seizure type and inhomogeneous recording channels (both for number and type), it's not possible, for now, to build a model able to predict the arrival of a generic epileptic attack with reasonable temporal precision.

Contents

Chapter 1

Introduction

1.1 What are the epileptic seizures and epilepsy disease?

The definitions of **seizure** and **epilepsy** have changed during the time. The first meaning of seizure was "take hold" and derives from the Greek culture.
In modern popular language, instead, people uses the word *seizure* to describe a sudden and severe event (for example, “he had a heart seizure").
However in order to define many physical or psychological sudden events related to epilepsy disease, the International League Against Epilepsy and the International Bureau for Epilepsy have come to consensus definitions for both *epileptic seizure* and the *epilepsy* [30]:

Epileptic Seizures	transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain.
Epilepsy	disorder of the brain characterized by an enduring predisposition to generate epileptic seizures and by the neurobiologic, cognitive, psychological and social consequences of this condition. The definition of epilepsy requires the occurrence of at least one epileptic seizure.

Table 1.1: Definitions of epilepsy and epileptic seizures

1.2 Epileptic seizures

Although clinical symptoms linked to an epileptic event can never been immediately recognized, it’s evident that each ES has a clear beginning and end.
Using the EEG and patient’s behavioral clinical analysis the physician could be able to determine the time interval when ES happens, as shown in figure 1.1 .

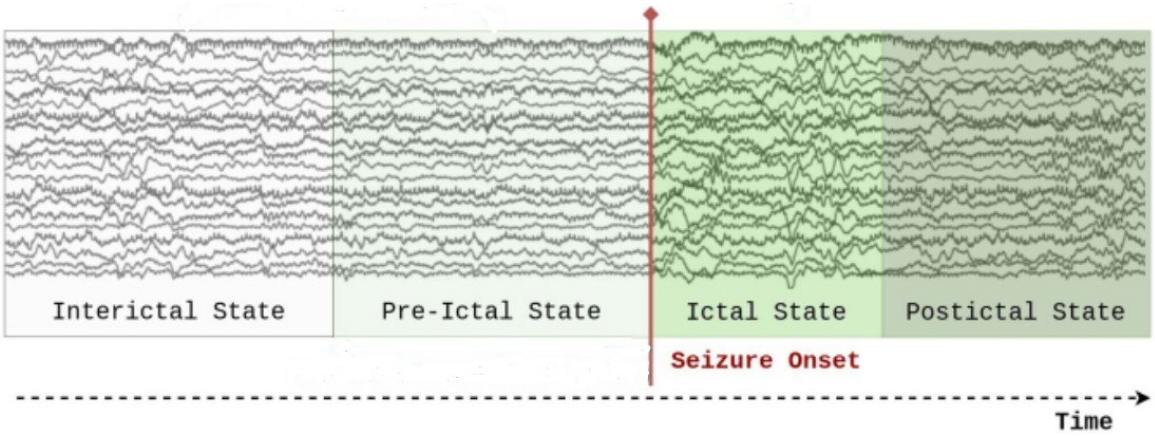


Figure 1.1: EEG seizure recording: the **ictal state** is the seizure's time window

However, discrepancies between what doctors see and what medical devices measure lead to inaccurate time estimates.
In reality, many aspects influence the presentation of seizures: location of onset in the brain, propagation patterns, brain maturity, confounding disease processes, sleep-wake cycle, medications, etc.

For example, it is commonly used to see a wide repertoire of jerky movements in preterm infants, many of which are essential for normal sensory-motor development.
It can be difficult to distinguish these essential jerky movements from seizures; this often leads to unnecessary treatment of infants with antiepileptic drugs [27].

On the other hand, some EEG patterns are neither perceived by the patient nor by the observer (sometimes incorrectly called *subclinical seizures*) and therefore cannot be properly defined as epileptic seizures [30].
Therefore a simultaneous medical examination performed by directly checking the patient reactions and the medical instrumentation could lead to a correct estimation of when and which types of ES occurs.

1.2.1 Classification

In the last classification [31], seizures are classified according to three characteristics:

- Origin of the seizure in the brain: focal or generalized.
- Degree of awareness during the seizure: intact or impaired awareness.
- Level of body movement: motor or non-motor

According to these features, ES are classified into three macro-groups:

1. **Focal onset seizures**, also subsided into three subtypes:
   (a) Retained/impaired awareness.
   (b) Motor/non-motor onset.
   (c) Focal to bilateral tonic-clonic.
2. **Generalized onset seizures**: motor or non-motor (absence) onset seizures.
3. **Unknown onset seizures**: motor/non-motor or unclassified [5].

1.2.2 Focal onset seizure

Focal onset seizure or a focal seizure is when ES starts in one hemisphere of the brain and progresses into the other. Nowadays these seizures are also be known as **partial seizures**.
The awareness in focal seizure varies and can be: present or absent.
If a seizure occurs when the patient is fully aware of the events occurring around him it's called as *focal aware seizures*.

However, in most cases, the patient may lose awareness of some events; when this occurs, a focal *impaired awarness* takes place.
This type of event may be associated with loss of memory (*amnesia*) during seizures [11].
Classification based on awareness is fully optional and can be omitted, especially when it is difficult to establish the level of awareness impairment as in sudden and brief ES [17].

On a general basis, before a seizure event happens, the patients experience a several feelings known as *aura*: deja vu, a strange taste or smell, a rising sensation in the stomach, lip-smacking and hand rubbing. In this case, the aura is also known as a focal aware seizure [11].
Putting a focal seizure in a category is used primarily to take into account the first signs or symptoms that appear at the beginning of the ES event (**the Rule of First**).
An exception occurs when it deals with *impairment of awareness* (IOA).
Only in this scenario doctors always define an ES as an IOA seizure although symptoms occur in the last phase of the event.

The Rule of First is also useful to determine the area of the seizure [16, 14]: for example, a focal seizure could be associated with jerking of a single limb (arm or leg), which can progress to involve both sides of the limbs (both arms and legs), where it is referred to as a focal to a bilateral tonic-clonic seizure.

1.2.3 Generalized seizures

Generalized seizures are defined as abnormal electrical activities that start in both the right and left cerebral hemispheres and then spread to the other brain neuronal networks [11].
They could be divided into motor or non-motor classes:

- *motor seizures* include tonic-clonic, clonic, tonic, myoclonic, myoclonic-tonic-clonic, myoclonic-atonic, atonic, or epileptic spasms.

- *no motor seizures* are typical or atypical absence seizures or seizures with myoclonic activity or eyelid myoclonia.

1.2.4 Unknown seizures

Seizures of unknown onset are the ones that occur either when the patient is asleep or has not the required awareness to describe them.
Unclassified seizures also occur when the clinician is certain about information concerning the seizure event but it's not able to describe it due to incomplete information he has on it.
Therefore, the term unknown onset is just like a nickname but not the characteristic of the seizure [12].

1.3 Epilepsy disease

1.3.1 Epidemiological data overview

The World Health Organization (WHO) and its partners have recognized epilepsy as a major public health concern affecting people of all ages, races, social classes and geographical locations [6].
Around fifty million people worldwide are affected by epilepsy, making it one of the most common neurological globally diseases [4].

1.3.2 Incidence

In a systematic review and meta-analysis of incidence studies, the pooled[1] incidence rate[2] of epilepsy defined as $\frac{n\ cases\ during\ observation\ period}{total\ persons-time\ of\ observation\ while\ at\ risk\ during\ study}$ was 61.4 per 100,000 person-years (95 % $confidence\ interval$ (CI) 50.7–74.4) [15].
The incidence was higher in low/middle-income countries (LMIC) than in high-income countries (HIC), 139.0 (95 % CI 69.4–278.2) vs 48.9 (95 % CI 39.0–61.1) .

This can be explained by the different structure of the risk populations and increased exposure to perinatal risk factors, higher rates of central nervous system (CNS) infections and traumatic brain injury (TBI) in LMIC.
The incidence of epilepsy is also higher in the lowest socioeconomic classes in HIC and, within the same population, for people of different ethnic origin [8].

1.3.3 Prevalence

The prevalence[3] of the epilepsy is defined as $\frac{patients\ at\ time\ t}{population\ at\ time\ t} \times K$ where K is a constant to improve readability and it differs significantly among countries depending on the local distribution of risk and etiologic factors, the number of seizures at diagnosis and if considering only active epilepsy (active prevalence) or including also cases in remission.

In Fiest et al. [15], the overall lifetime prevalence (the proportion of individuals in a population who at some point in their life (up to the time of evaluation) have experienced a 'case') of epilepsy was 7.60 per 1,000 people (95 % CI 6.17–9.38) and was higher in LMIC (8.75 per 1,000; 95 % CI 7.23–10.59) than in HIC (5.18 per 1,000; 95 % CI 3.75–7.15).
The point prevalence of active epilepsy was 6.38 per 1,000 (95 % CI 5.57–7.30).
The median point prevalence of active epilepsy in LMIC was 6.68 (95 % CI 5.45–8.10) and in HIC was 5.49 (4.16–7.26).

[1] Pooled analysis : "A statistical technique for combining the results of multiple epidemiological studies when individual studies are too small to allow any definite conclusion." $http://icrpaedia.org/Pooled\ analysis$
[2] https://sphweb.bumc.bu.edu/otlt/MPH-Modules/PH717-QuantCore/PH717-Module3-Frequency-Association/PH717-Module3-Frequency-Association4.html
[3] $https://it.wikipedia.org/wiki/Prevalenza$

Prevalence estimates also vary and tend to be higher in individuals of certain ethnicities [1], people with poor health and socially deprived subjects [20].
Along with issues in the study design, the demographic structure of the study population, the prevalence of environmental risk factors, the quality of health management can be implicated [6].

1.3.4 Prognosis

Looking at the statistics of epileptic patients, up to 80 % who get medical therapies enter into prolonged periods of seizure remission and up to 50 % continue to be seizure-free after discontinuation of treatment [3, 36].
However, reports from several LMIC (where treatment gap is high) give prevalence and remission rates overlapping to HIC [7].

The risk of relapse after a first unprovoked seizure in some studies was fairly consistent with the rates of 36 to 37 % at 1 year and the rates of 43 to 45 % at 2 years [36] and in a more systematic review, the mean risk of recurrence was 51 % (95 % CI 49–53 %) [9].
According to some studies [9], the probability of relapse, after an unprovoked seizure, decreases with time;in particular about 50 % of recurrences fall within the following 6 months.

Taking into account two of the main predictors of recurrences, the etiology of the seizure and an abnormal (epileptiform and/or slow) EEG, it's possible to make a sort of *risk list* about pooled 2-year recurrences:

Lowest	for an idiopathic or cryptogenic first seizure with a normal EEG (24 %; 95 % CI 19–29 %)
Intermediate	for a remote symptomatic seizure (48 %; 95 % CI 34–62 %) with normal EEG or an idiopathic/cryptogenic seizure with an abnormal EEG (48 %; 95 % CI 40–55 %)
Highest	with a remote symptomatic seizure with an abnormal EEG (65 %; 95 % CI 55–76 %)

Table 1.2: Prognostic risk levels

Other potential risks of recurrence of seizures come from several factors such as interictal EEG epileptiform abnormalities, seizures occurring during sleep, focal seizures, family history, etc.
LMIC verified the hypothesis that spontaneous remission of epilepsy is a common event where epsilepsy is largely untreated (range between 70-94 %) [26].

For example, in Ecuador the cumulative annual incidence rate was 190 per 100,000 and the prevalence rate of active epilepsy was 7 per 1,000, which implies a remission rate of at least 50 % [29].

However as proposed by Sander [10], epilepsy patients can be classified into 4 different prognostic groups:

Excellent	(about 20–30 % of the total) High probability of spontaneous remission; these include benign focal epilepsies, benign myoclonic epilepsy in infancy and epilepsies provoked by specific modes of activation, that is, reflex epilepsies.
Good	(about 30–40 %) Easy pharmacological control and possibility of spontaneous remission; these include childhood absence epilepsy and some focal epilepsies.
Uncertain	(about 10–20 %) Patients respond to drugs but tend to relapse after treatment withdrawal; these include juvenile myoclonic epilepsy and most focal epilepsies.
Poor	(about 20%) Seizures tend to recur despite intensive treatment; these include epilepsies associated with congenital neurological defects, progressive neurological disorders and some symptomatic or cryptogenic partial epilepsies.

Table 1.3: Prognostic seizures

1.3.5 Mortality

As prevalence and incidence also, mortality depends on the precision of the information on causes of death and the survey methods [39] . In general, epilepsy has a low mortality risk, but differences occur taking into account factors such as age, type of seizures, incidence, prevalence, etc. [40].
In LMIC, the lack of access to health facilities, lead the mortality ratio (MR) to 19.8 % (95 % CI 9.7-45.1) almost 10 times higher then HIC where MR range between 1.6 and 3.0 [40, 22]. The MR is also slightly higher in men than in women and in children and adolescents.

The sudden unexpected death in epilepsy (SUDEP) is the most important cause of death due to epilepsy or ES.
The incidence of SUDEP among people with epilepsy is 1.2 per 1,000 person-years (95% CI 0.9–1.5) and ranges from 1.1 (95% CI 0.5–2.3) in children under age 16 years to 1.3 (95% CI 0.9–1.8) in adults after the age of 50 years [37].
In particular the major risk factors include generalized tonic-clonic seizures, nocturnal seizures, and persistence of seizures [18].

1.3.6 Classification types

Today, the classification framework for epilepsy takes into account a multidimensional framework in which both the type of seizures and the etiology are needed to correctly categorize [31, 35].

FOCUS ON SEIZURES	
Focal Epilepsy	is a continuum of seizures (unifocal and multifocal disorders) that are confined to one hemisphere and can present with any type of focal seizure. It is diagnosed clinically but supported by an interictal EEG with focal epileptiform discharges.
Generalized Epilepsy	consists of many seizure types and a patient can present with any type of generalized seizure: either motor or non-motor. Diagnosis is based on generalized spike-wave activity on EEG and clinical presentation.
Combined generalized and focal Epilepsy	is a condition in which a patient has both focal and generalized seizures (which is usually diagnosed by EEG) and commonly occur in infants or children with severe epilepsies.
Unknown Epilepsy	is when the clinician cannot categorize seizures as either focal or generalized, especially in a resource limiting setting.

Table 1.4: Epilepsy based on seizures

FOCUS ON ETIOLOGY	
Idiopathic Epilepsy	epilepsy of predominately or presumed genetic origin such that there is no gross neuroanatomic or neuropathologic abnormality.
Symptomatic Epilepsy	epilepsy of an acquired or genetic cause, associated with gross anatomic or pathologic abnormalities (and/or clinical features) indicative of underlying disease or condition.
Provoked Epilepsy	epilepsy in which a specific systemic or environmental factor is the predominant cause of the seizures such that there are no gross causative neuroanatomic or neuropathologic changes.
Cryptogenic Epilepsy	epilepsy of presumed symptomatic nature in which the cause has not been identified. The number of such cases is diminishing, but currently this is still an important category, accounting for at least 40 % of adult-onset cases of epilepsy.

Table 1.5: Epilepsy based on etiology[4]

[4] A more recent epilepsy etiology classification was proposed by Fisher et al [14].

Anyway, different classifications exist and take into account more factors: Luders and colleagues, for example, proposed one composed of clinical semiology, location of the disease (epileptogenic zone), etiology and comorbidity.

This classification is called the Four-Dimensional Epilepsy Classification System [25] and also used a nonspecific terminology like *paroxysmal events* in order to give a definitive diagnosis. The more complex scheme is offered by Loddenkemper et al [23]: it includes the aforementioned four dimensions with seizure frequency as the fifth dimension. Seizure frequency refers to the number of episodes of seizure events that occur over a given period of time:

1. *Daily*: when the seizure event occurs every day.

2. *Persistent*: when it occurs at least once in six months but not daily.

3. *Rare*: when it occurs at more than six months interval.

4. *Undefined*: when the frequency cannot be predicted. It includes breakthrough seizures that occur in a well-controlled patient due to either trigger like sleep deprivation or abrupt stopping of medications [23].

1.4 Why categorize?

The classification of epilepsy and seizures gives an important advantage both for patients and clinicians. By specifying each characteristic,the physicians are able to distinguish medications for every type of disease and helps them choosing specific drugs for each case [32].

From a pharmacological point of view, this classification has great relevance, since focal seizures, regardless of the part of the brain involved, were shown to respond well to a specific set of anticonvulsants. On the contrary, drugs for generalized seizures depend on their specific type [24].

1.5 Related works

This study is part of a current of research that in recent years is rapidly gaining ground and seeks to respond to the growing needs to predict the occurrence of ES with the aid of artificial intelligence, in particular, with deep learning models.

In fact, although several studies have only made use of feature extraction methods such as the Support Vector Machine (SVM), Random Forest, K-nearest neighbors as in the work of Resmi Cherian, E. Gracemary Kanaga [13] or as in the work of Yanli Yang et al. [41] it has been observed that the performance obtained is generally lower than those related to deep learning architectures, even if the latter require a larger amount of data to be trained correctly [34].

For example, working with preterm EEG infant, the article written by O'Shea et al. [27] shows how an alternative deep learning approach has a more stable trend then SVM when it tested on the preterm cohort, starting with an area under the ROC curve (AUC) of 93.3% for the term-trained algorithm and reaching 95.0% by transfer learning from the term model using available preterm data.

Ranjan Jana and Imon Mukherjee [19] propose an efficient method of predicting seizures using raw EEG signals linked to a channel reduction technique.
An accuracy ($\frac{correct\ classifications}{all\ classifications} \times 100\%$) of 99.47 % has been achieved for the classification of preictal and interictal states and the method can predict seizures before 10 min with a sensitivity ($\frac{True\ positive}{True\ positive + False\ Negative} \times 100\%$) of 97.83 % and a specificity ($\frac{True\ negative}{True\ negative + False\ Positive} \times 100\%$) of 92.36 %.
But this proposed method implemented a patient-specific seizure prediction and cannot predict seizures efficiently for other epilepsy patients.

Also Yankun Xu at al. [19] propose an end-to-end deep learning solution using convolutional neural network (CNN) to classify interictal and preictal states: one and two dimensional kernels are adopted in the early- and late-stage convolution and max-pooling layers,respectively. The proposed CNN model is evaluated on Kaggle intracranial and CHB[5]-MIT[6] scalp EEG datasets. The overall sensitivity, false prediction rate and AUC reach 93.5%, 0.063/h, 0.981 and 98.8%, 0.074/h, 0.988 in two datasets respectively.

Applying the ResNet-50, AlexNet and VGG networks (advanced CNN models with different layers) the study conducted by Mehmet Akif Ozdemir, Ozlem Karabiber Cura and Aydin Akan [33] demonstrates how to input frequency-time images, it is possible to arrive at a binary classification between *inter-seizure* and *pre-seizure* with an accuracy of 98 %.

1.6 Work's aim

Using spectrograms, this work aims to classify the arrival of a seizure up to 2 hours by classifying each EEG section into tranches of 30 minutes each.

The challenge is particularly demanding as we do not start from pre-packaged datasets such as CHB-MIT or made by type of seizure/patient's age (as was done in the previous articles), but we start from EEG of patients with different ages, unknown suffered seizure and a different number of channels in each recording.

Actually, the research point to generalize the learning models to real cases where by play-force there is not a priori information about patient.

[5] Boston Hospital Center
[6] Massachusetts Istitute of Technology

Chapter 2

Material and methods

2.1 Dataset

The dataset is provided by the Micromed company and is composed of two tranches of patients of 20 and 18 units, respectively.
The EEG recordings in each patient refer to different time periods and for practical reasons each recording session is divided into files whose maximum size is equal to 2 Gb.
The Brainstorm[1] software written in Matlab language was used to view the data contained in the individual TRC files.

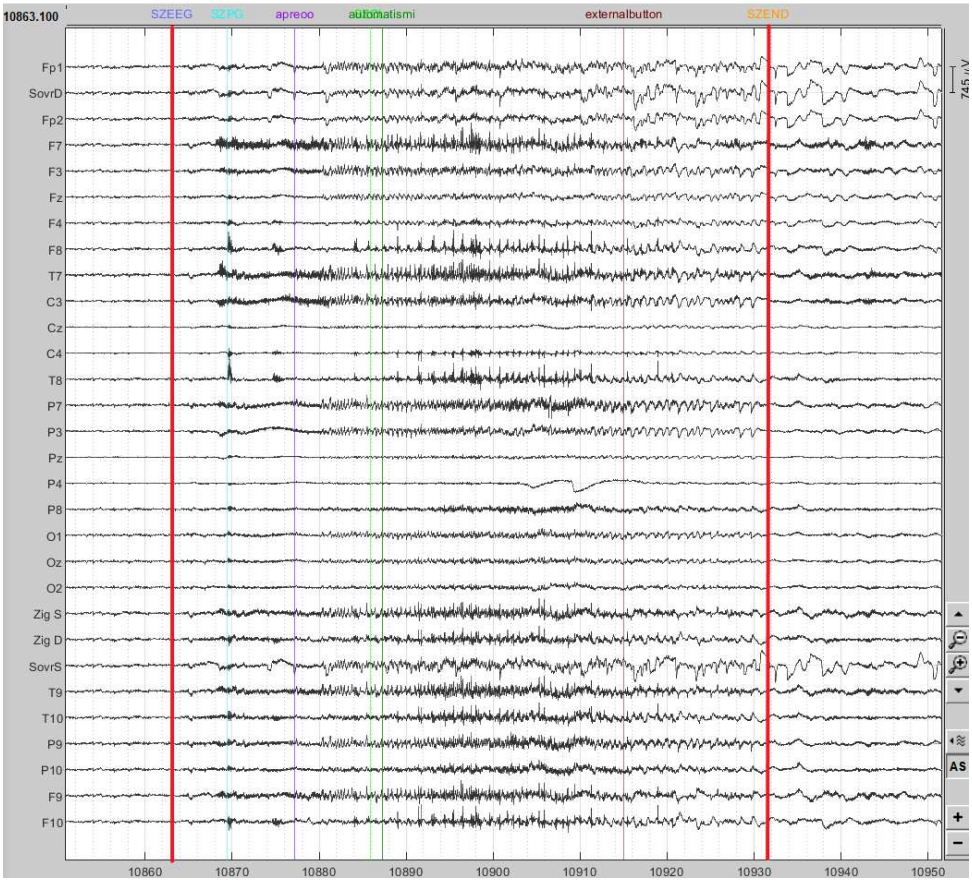


Figure 2.1: Example of Brainstorm EEG visualization with channels

[1] https://neuroimage.usc.edu/brainstorm/Tutorials

As shown in the figure 2.1 once the connection with the raw file has been created, it is possible to view the trend of the signals in the respective channels (shown on the left) as a function of time.
In addition, Brainstorm also allows you to label each event and represent it on the (s, μV) graph: in the previous example, the focus was on the beginning and the end of the ictal phase.

2.2 Preprocessing

It started with the extraction of the events: using the Brainstorm interface, it is possible to obtain the events related to the single examined TRC file and the instant in time in which they occur.
It was realized that the same event can be described with different names: for example, the end of the seizure is sometimes referred to as EZEND/ezend, other times as EGEND and rarely presents forms of the SZEEGEND type.

Therefore, it was necessary to replace the different labels with a single one capable of uniquely defining the event without ambiguity: for our purposes, it was essential to establish EZEEG as the beginning of the seizure and EZEND to indicate its end.
In this way it was easier to program a Python script able to select only the events useful to us and consequently discard the others.

Only the data files containing at least one seizure were taken into account: the goal of this research is to predict an epileptic seizure with a maximum of 2 hours in advance and as each file has an average of 8 hours of data taking it's not useful to hold also back the remaining files without significant events.
In this first selection, attention was paid to ensure that each seizure was truly independent from the others, setting a temporal distance of at least 4h from each other.

The effects of *post ictal phase*, indeed, can also spread for a long time after the EZEND event, and therefore it is necessary to be sure that you are in a section of the EEG that is actually not related to the others.
Among the different patients and for the same patient, there was a different number of channels used to record brain activity and also a different sampling frequency of the signals.

Having found that the range of channels varies from a few tens to hundreds of units, it was decided to standardize the dataset by taking only those files whose channels vary in number from 24 to 35: in this range, very noisy channels are also excluded because not provide useful information about features extraction.

This second selection was a choice determined by the desire not to exclude too many files and, at the same time, to avoid such a marked difference between them that would have had effects on the efficiency of the scripts.

After these two selections step, the dataset counts 32 effectively patients.
In this range of channels, a subset of 16 channels has been found to be common in all recordings and are symmetrically distributed in the front, middle, and back of the brain.

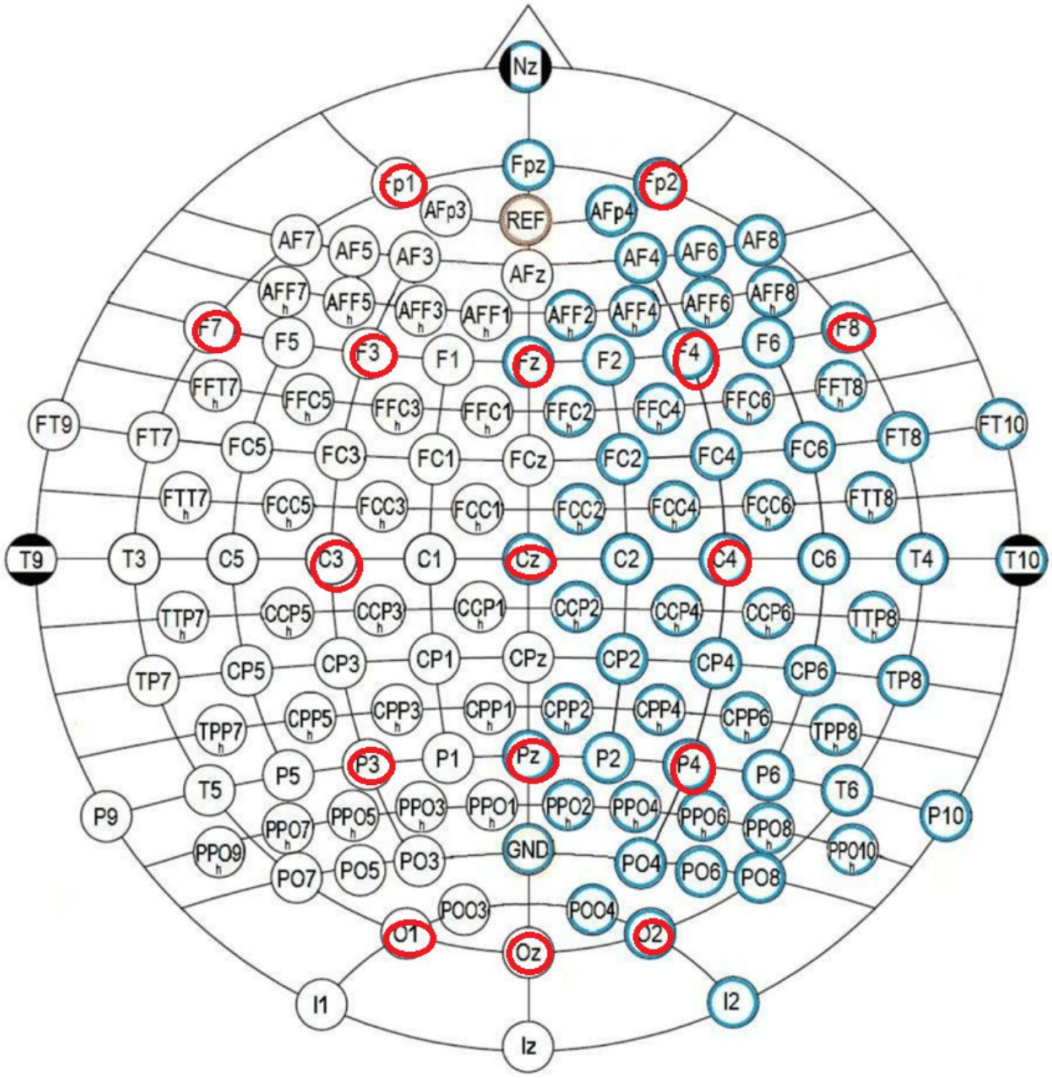


Figure 2.2: Common 16 channels

Once the dataset is ready,the following pipeline starts:

1. Each TRC file is read using the neo.io.Micromed package: both the μV signals relating to the different channels and the relative sampling frequency are extracted.

2. Once the dataframe has been built with μV values as rows and channels as columns, many other columns (with null signal) are added until their number is exactly 35.
   This is fundamental for CNN training because it will give it the ability to feed it with arrays of equal size.

3. Only parts of the dataframe related to the *no-ictal* EEG are analyzed and their temporal order is reversed so that the first lines are those closest to the seizure. These parts are then divided by a length corresponding to 5 seconds.

4. These parts, one after the other, undergo into a process of downsampling to 256 Hz. However, for each channel, using the *spectrogram* function[2] (provided by the *scipy* library) the relative spectrogram is created and the amplitudes are modulated on a logarithmic basis.The sampled frequencies undergo a system of filters to eliminate those components of the spectrum that would otherwise *dirty* the signal [2, 38, 28].
So an high pass filter at 1 Hz, a low pass filter at 125 Hz and a notch filter at 50 and 100 Hz are implemented.

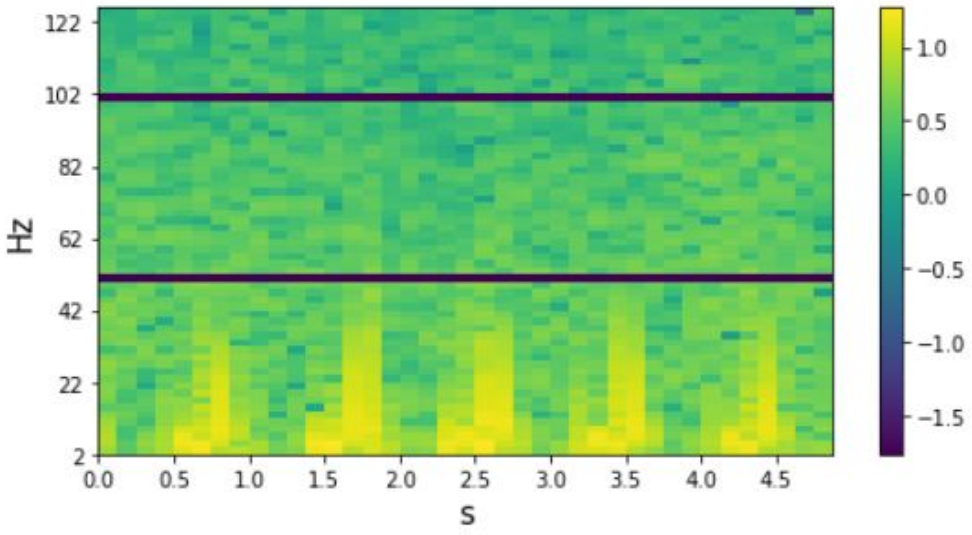


Figure 2.3: spectrogram visualization

5. Finally, each *packet* of 35 spectra (one per channel) is labeled according to the temporal category to which it belongs. Considering the beginning of the seizure part is defined by EZEEG event, the labels are:

   - "0" if the recording part falls between 0 s and 1800 s

   - "1" between 1800 s and 3600 s

   - "2" between 3600 s and 5400 s

   - "3" between 5400 s and 7200 s

2.3 Processing

At this point we proceed with the processing of the dataset, evaluating how balanced it is between the various categories.

[2] https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.spectrogram.html

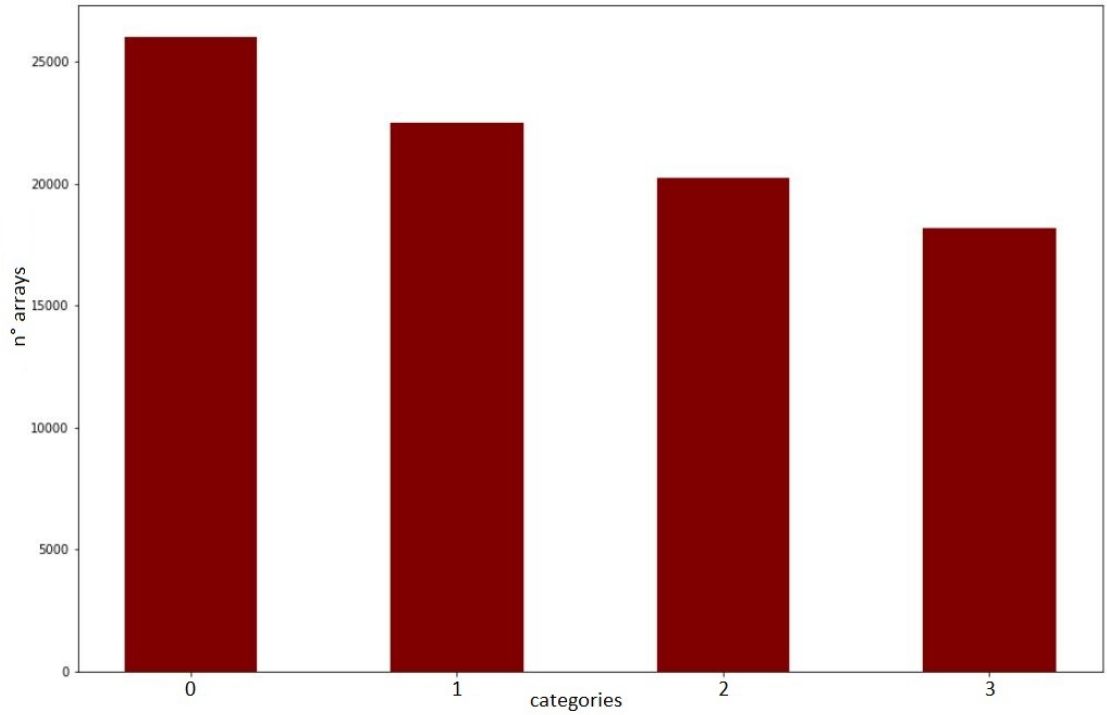


Figure 2.4: Unbalanced dataset

Clearly, as was easy to guess, the "0" category is more populated than the others because being the temporal section immediately before the epileptic event is certainly present in every TRC while the same cannot be said for the more distant categories. So, first it's mandatory to balance the dataset.

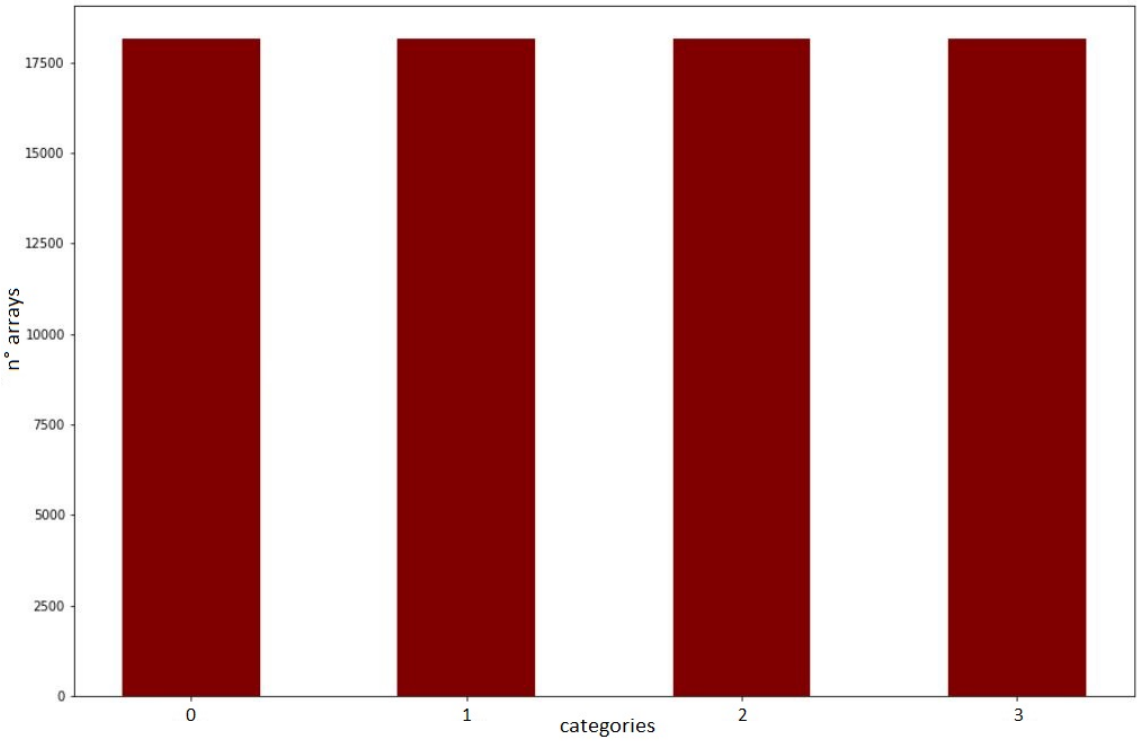


Figure 2.5: Balanced dataset

At this point the balanced dataset is divided into a training part and a test part, respectively 90 % and 10 % of the total. To obtain this splitting, the array-packets (with their respective labels) are first mixed and then randomly selected so that the size of both parts is exactly that indicated by the percentages. Since the choice is made randomly, the 10 % test part will be called as **random test** .

As a matter of optimization of the classification algorithm, the mean and standard deviation of the training part are computed and these two parameters are used to normalize the spectra of both the training and the random test set.

Training dataset
$\mu = 87.16264$
$\sigma = 44.10327$

Table 2.1: Mean and std from training dataset

2.4 Model

2.4.1 Why CNN and how it works

Since the dataset is now composed by spectrograms (i.e. *images*) it was decided to build a two-dimensional convolutional network model known as CNN.
The idea is to train an architecture able to extract the intrinsic features in the training images, *memorizing* the patterns found and subsequently recognizing them in new inputs during the model testing phase.

In order to do this, CNN has internal filters whose weights are calibrated during the training phase thanks to a process called backpropagation error method (BEM).
In practice, CNN is connected to another neural network called fully connected network (FCN) which in its last layer has a number of outputs equal to the number of categories.

During the training phase, for each input (batch) provided, the model tries to hypothesize the correct category to which it belongs and to evaluate the error made during this attempt.
The error also depends on the loss function chosen during the compilation phase of the model.
Regardless of the loss function chosen, CNN + FCN, by means of BEM, updates the weights contained in the convolutional filters and in the FCN layers in order to minimize the error made on the training set.

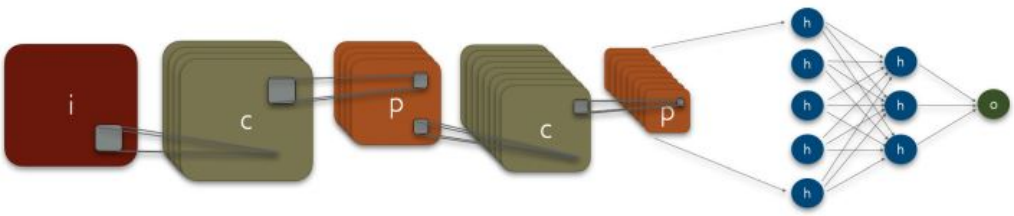


Figure 2.6: Example of CNN + FCN

To explain better BEM let be J(θ) the sum of the loss function terms $L(x^i, y^i, \theta)$ used: it depends by the general weights θ, the *input* x^i and the corrisponding *output* y^i with i goes from 1 to m, the number of training examples. In formula:

$$J(\theta) = \sum_{i=1}^{m} \frac{L(x^i, y^i, \theta)}{m} \tag{2.1}$$

First the loss gradient is computed with respect to weight:

$$\nabla_\theta J(\theta) = \sum_{i=1}^{m} \frac{\nabla_\theta L(x^i, y^i, \theta)}{m} \tag{2.2}$$

then the weights are updated:

$$\theta \leftarrow \theta - \eta \nabla_\theta J(\theta) \tag{2.3}$$

where the η is the learning rate. But using all the training set the algorithm becomes too slow, so typically a part of the samples n $<$ m is used:

$$g = \nabla_\theta J(\theta) = \sum_{i=1}^{n} \frac{\nabla_\theta L(x^i, y^i, \theta)}{n} \tag{2.4}$$

$$\theta \leftarrow \theta - \eta g \tag{2.5}$$

These last two methods are respectivelly called Gradient Discend (GD) and Stochastic Gradient Discent (SGD) and they are the keystones on which BEM works. Although the SGD aims to reach that the minimum of the loss function, in the most of the cases the parameter landascape is just locally convex :

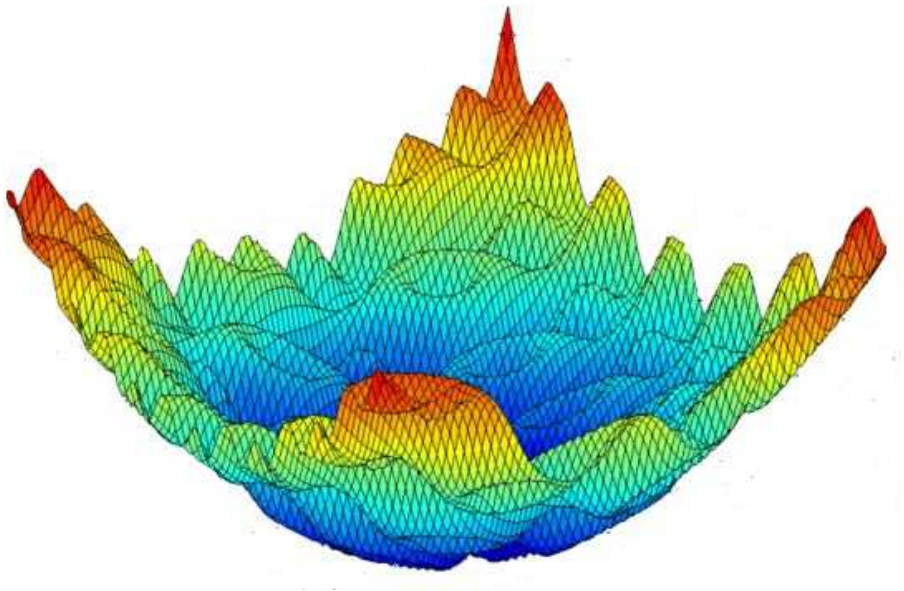


Figure 2.7: Non convex parameters landscape

So in order to avoid that a local minimum is reached and trapped the loss descent, some advance optimizers are developed: AdaGrad,SGDNesterov,AdaDelta,RMSProp and Adam[3].

[3] https://elearning.dei.unipd.it/pluginfile.php/627867/modresource/content/1/NNDL06AdvancedOptimzation.pdf

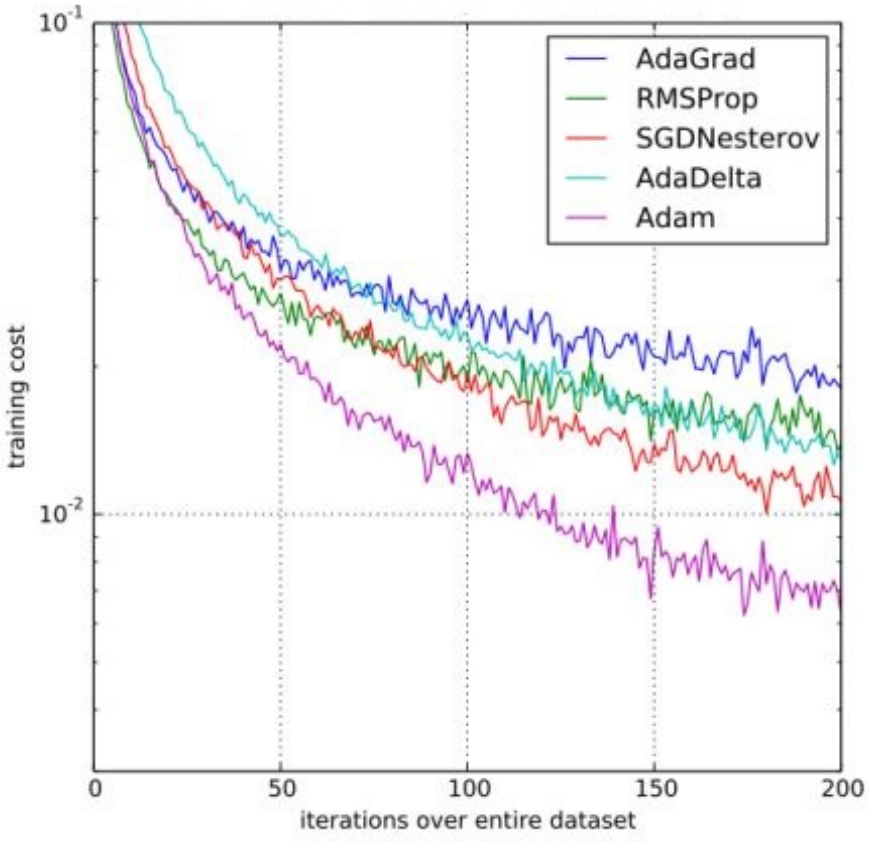


Figure 2.8: Optimizers performance on MNIST dataset

As shown above, Adam has the better performance then the others and it has also a very poor training cost. In formulae Adam could be express as:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{2.6}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \tag{2.7}$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{2.8}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{2.9}$$

$$w_t = w_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{2.10}$$

where g_t and g_t^2 are the first and the second moments of the gradients, $\hat{m}_t$ and $\hat{v}_t$ are the correct estimates (unbiased) of m_t and v_t instead ϵ,β_1 and β_2 are useful parameters to properly define the method .

As well as decreasing the error and reach the loss function global minimum, it's also important to set the best possible model configuration tuning the so called hyperparameters.

2.4.2 What are the CNN hyperparameters and how tune them

By hyperparameter it generally meant a variable that is determined by the user and is therefore not intrinsically predetermined by the layer or by the model itself.
There are several types of hyperparameters, for example:

- *loss function* : the function to be minimized during the model training phase.

Loss function name	Formula	Explanation
Multi-Class Cross Entropy	$-\sum_{c=1}^{M} y_{o,c} log(p_{o,c})$	• M = number of the classes. • y = binary indicator (0 or 1) if class label c is the correct classification for observation o . • p = predicted probability observation o is of class c .
Binary Cross Entropy	$-$ylog(p) + (y$-$1)log(1$-$p)	as above but M = 2
RMSE	$\sqrt{\frac{\sum_{c=1}^{m}(h(x^i)-y^i)^2}{m}}$	• m = number of samples. • x^i = i-th sample from dataset. • $h(x^i)$ = prediction for i-th sample (thesis). • y^i = ground truth label for i-th sample.
MSE	$\frac{\sum_{c=1}^{m}(\hat{y}^i-y^i)^2}{m}$	as RMSE but $h(x^i) \rightarrow \hat{y}^i$

Table 2.2: Examples loss functions

- *n°epochs* : how many times the model has to train on the entire dataset.

- *batch size* : how many inputs the model receives at a time.

- *learning rate* : is a parameter of the optimizer that establishes the *step* used at each iteration to settle the weights to achieve the global/local minimum of the loss function.

- *filter size* : the size of the kernel's filters which in turn convolution on the input[4].

[4] https://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/

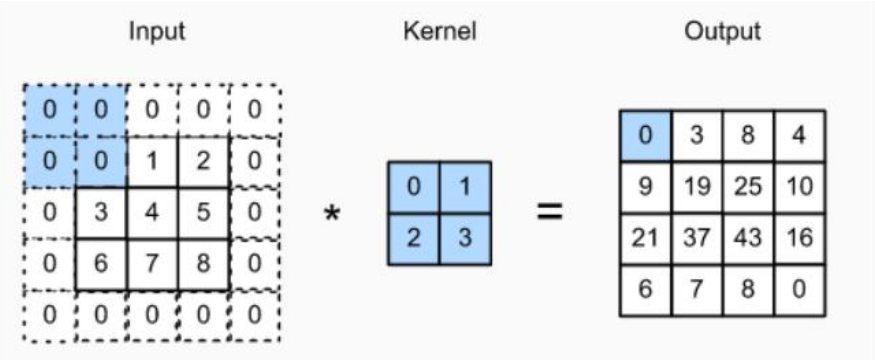


Figure 2.9: Kernel's filters convolution on the input

- *n°filters* : the number of filters to use.
- *stride* : how much the filter shifts to the input.

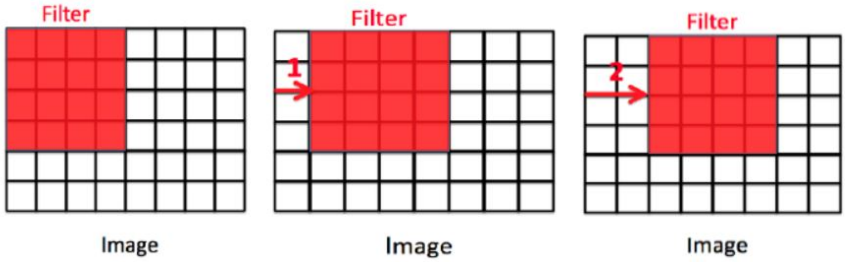


Figure 2.10: Stride

- *padding* : how to hem the input if the filter, moving on it, goes beyond the boundary.

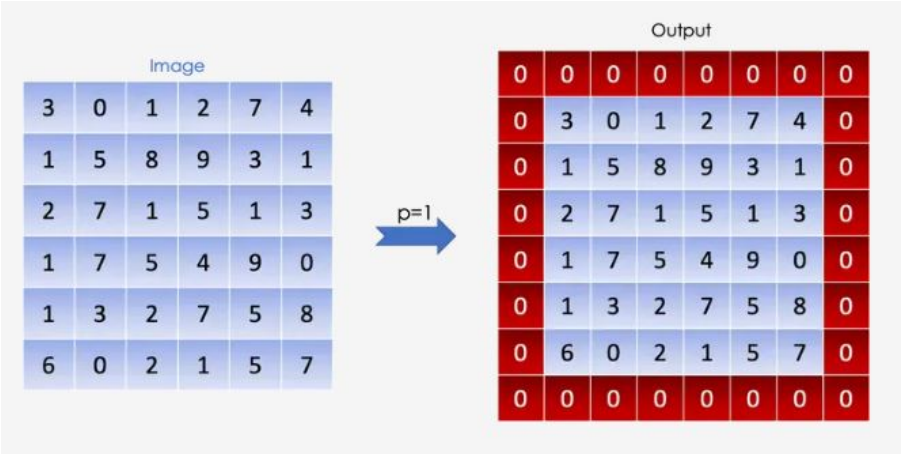


Figure 2.11: the new array has a frame made of zeros

- *activation function* : the function to apply to the layer before computing the final output.

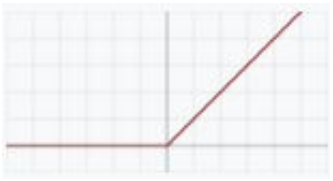
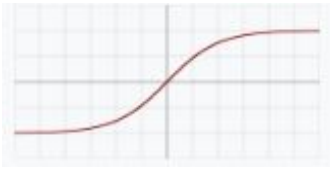
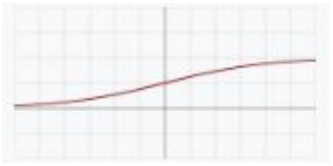
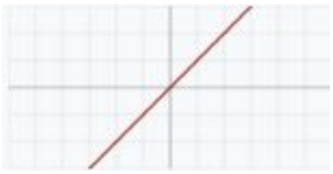
Activation function name	**Formula**	**Plot**
Rectified linear unit (ReLU)	$\max(0,x)$	
Hyperbolic tangent (tanh)	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$	
Logistic, sigmoid or soft step	$\frac{1}{1+e^{-x}}$	
Identity	x	

Table 2.3: Examples of activation functions

2.4.3 K-fold cross validation

A useful technique to configure these or other hyperparameters is certainly the k-fold cross validation. It consists of:

1. Fix a certain set of hyperparameters[5] that the user wants to evaluate with a given model architecture.

2. Choose a combination of hyperparameters.

3. Divide the training set into k-parts and one of these is defined as validation set while the remaining k-1 as actual training.

4. The model is trained on the k-1 parts and tested on the k-th.

5. At the end of the last epoch, points 1-4 are repeated until each part has been used as a validation set and the remaining one as training.

6. An average accuracy value is attributed to the just tested combination and the procedure restarts from point 2.

Finally, within the initial set, the combination that has reached the highest accuracy value is chosen and the model is calibrated accordingly.

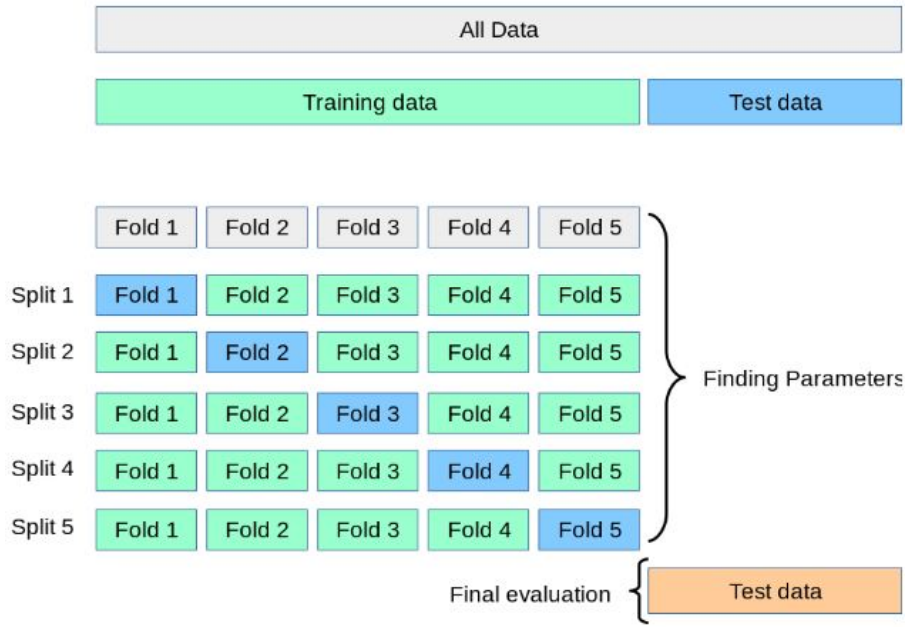


Figure 2.12: Example of k-fold cross validation

[5] Among them there can be one or more layers that can be added or removed

2.4.4 Architecture implemented

Using *tensorflow* library, an architecture based on three convolutional layers (CL) was built and each layer is followed by a maxpooling layer useful for optimizing the recognition of features and speeding up the training[6].

In particular 128, 256 and 512 are the number of the filters used for each CL respectively. Each filter size both in convolutional layer and max pooling one is 2×2 with unitary stride and no-padding. The activation function used by the CL is the ReLU cause of the non-saturation of its gradient, which greatly accelerates the convergence of SGD compared to the sigmoid / tanh functions [21].

These 6 layers are followed by another one which aim is to *flatten* the various arrays obtained in the three previous convolutions and make them suitable for classification.
The layer chosen (also by 3-fold cross-validation) is the GlobalAveragePooling2D (Gap) as it guarantees less information loss when passing from the matrix to the vector form[7].
Connected to this architecture there is a FCN with 4 output which has the purpose of classifying the input provided.

For this last layer the activation function is *softmax*: $\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$ where z is a vector of k real numbers (a sort of *sigmoid* genearalization) to give the probabilities for each class label.
Finally, the loss function used is *sparse categorical cross-entropy* (similar to *multi-class cross entropy*) because the classes are mutually exclusive.

Layer (type)	Output Shape	Param #
conv2d_99 (Conv2D)	(None, 61, 38, 128)	18048
max_pooling2d_99 (MaxPooling2D)	(None, 30, 19, 128)	0
conv2d_100 (Conv2D)	(None, 29, 18, 256)	131328
max_pooling2d_100 (MaxPooling2D)	(None, 14, 9, 256)	0
conv2d_101 (Conv2D)	(None, 13, 8, 512)	524800
max_pooling2d_101 (MaxPooling2D)	(None, 6, 4, 512)	0
global_average_pooling2d_16 (GlobalAveragePooling2D)	(None, 512)	0
dense_33 (Dense)	(None, 4)	2052

Figure 2.13: Architecture implemented and the training parameters for each layer involved.

[6] https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/
[7] https://tree.rocks/get-heatmap-from-cnn-convolution-neural-network-aka-grad-cam-222e08f57a34

To find the best possible configuration, three hyperparameters were tested: the learning rate η (related to the ADAM optimizer), Gap or Flatten layer (True or False in the next table) and the batch size.
Dividing the training set into 3 fold, cross-validation is used to test some combination of these 3 hyperparameters. The results are synthesized in the following table:

Gap	η	batch size	mean accuracy	mean epoch
True	10^{-4}	32	0.860604	97
True	10^{-4}	64	0.840338	98
True	10^{-4}	256	0.766205	95
True	5×10^{-4}	32	0.857468	92
True	5×10^{-4}	64	0.863877	86
True	5×10^{-4}	256	0.865406	96
True	10^{-3}	32	0.783977	77
True	10^{-3}	64	0.835077	96
True	10^{-3}	256	0.840675	94
True	5×10^{-3}	32	0.251659	17
True	5×10^{-3}	64	0.255330	3
True	5×10^{-3}	256	0.325945	10
True	10^{-2}	32	0.251659	3
True	10^{-2}	64	0.252898	2
True	10^{-2}	256	0.261448	1
False	10^{-4}	32	0.861185	88
False	10^{-4}	64	0.859899	99
False	10^{-4}	256	0.823663	99
False	5×10^{-4}	32	0.805280	39
False	5×10^{-4}	64	0.837809	55
False	5×10^{-4}	256	0.863000	73
False	10^{-3}	32	0.660035	32
False	10^{-3}	64	0.759891	20
False	10^{-3}	256	0.830850	98
False	5×10^{-3}	32	0.253953	4
False	5×10^{-3}	64	0.270595	2
False	5×10^{-3}	256	0.285574	1
False	10^{-2}	32	0.253953	1
False	10^{-2}	64	0.253877	3
False	10^{-2}	256	0.361374	94

Table 2.4: Cross validation results with mean number of epochs needed to achieve accuracy result

Below are the graphs related to accuracy and loss about the best configuration achieved: Gap = True, lr = 5×10^{-4} and batch size = 256.

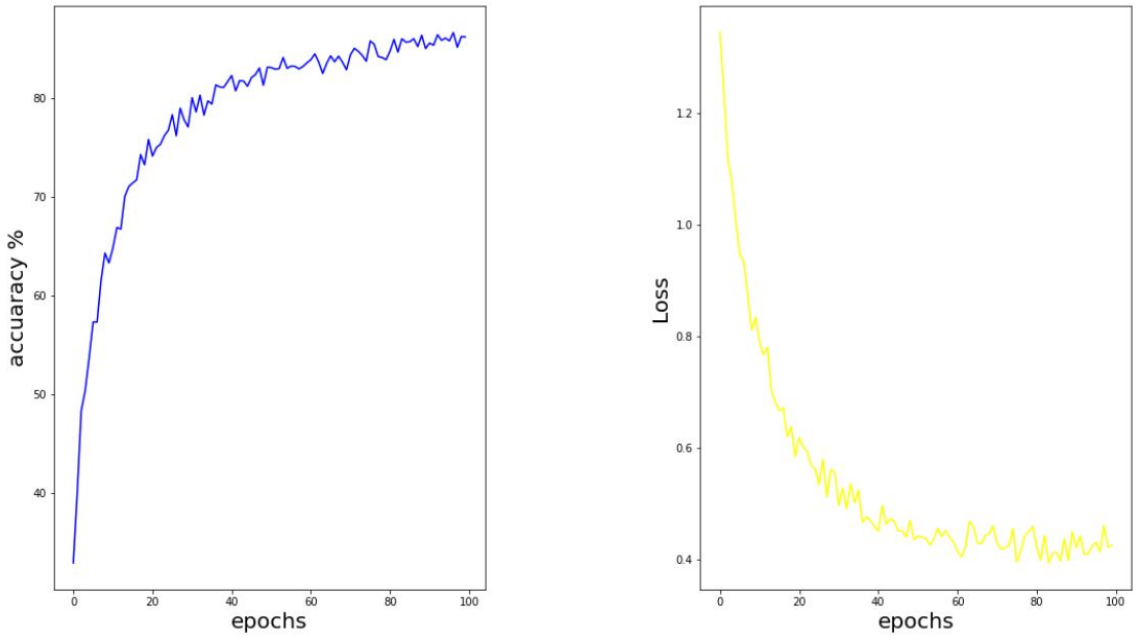


Figure 2.14: Accuracy trend (left panel) and Loss trend (right panel) for the CNN model as learning progresses.

Looking at the graphs above, it could be claimed that the model is *learning*: epoch by epoch the loss continuously drops and the accuracy increases.

Chapter 3

Results

By testing the architecture in figure 2.13 on the *random test*, it was obtained a result of approximately 88% of accuracy captured by the confusion matrix (CM) : each row of the matrix represents the instances in an actual class while each column represents the instances in a predicted class.

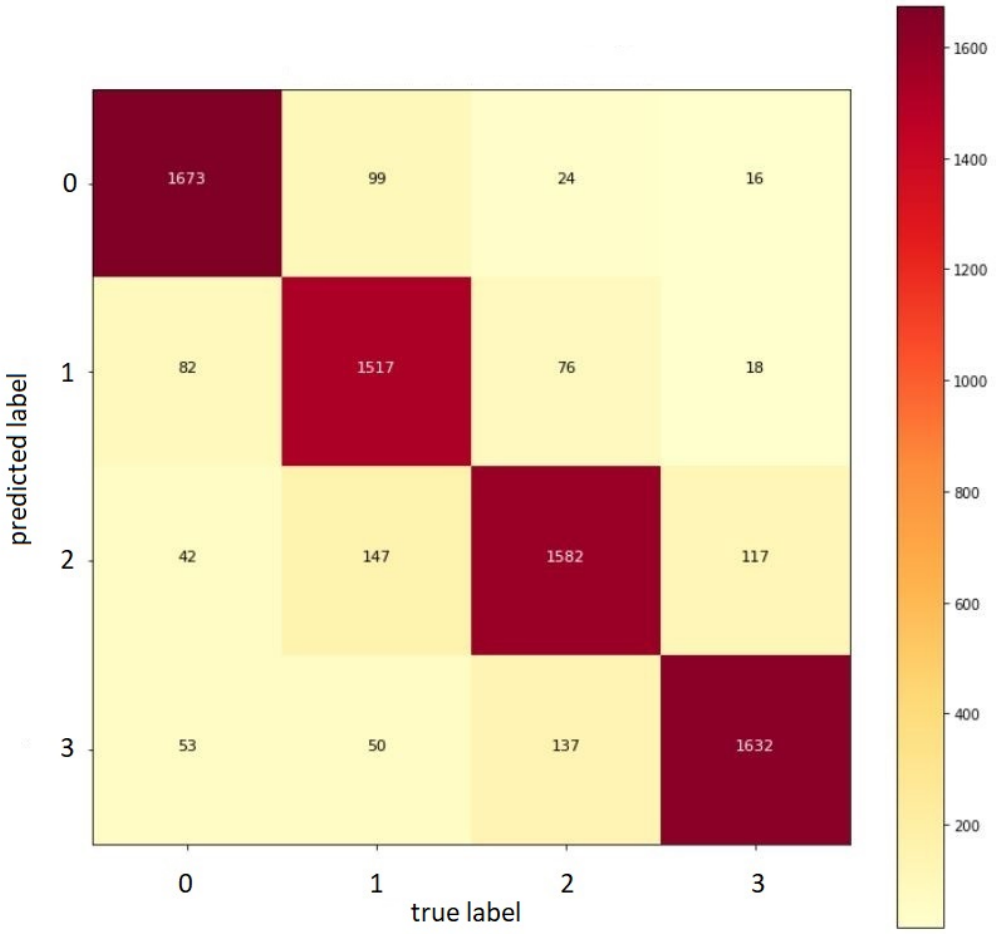


Figure 3.1: Confusion matrix for random test

However, if this new architecture is tested once again on new patients (in this case the last two) there is a lowering of performance up to 28%, that is, a little higher than pure chance.

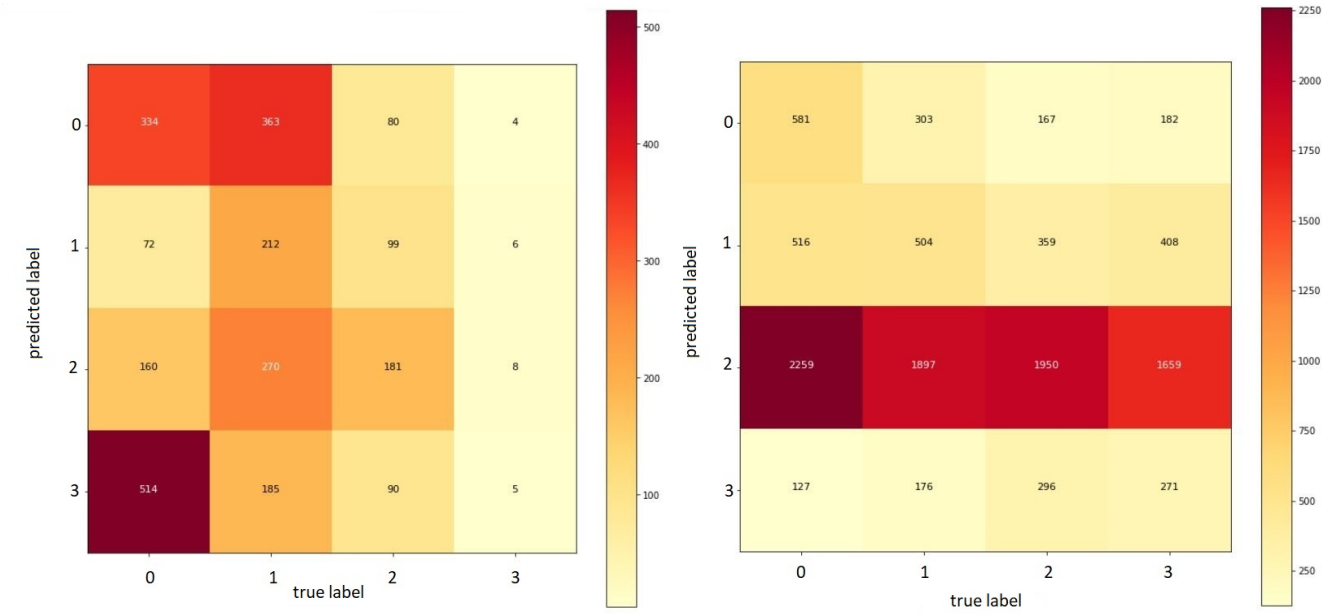


Figure 3.2: last two patients confusion matrix

As shown above, the last two CMs are quite different from each other: looking at the left one, a sort of slight decrease along the first diagonal should be noticed, but, on the contrary, on the right side it seems that only the third category is interested by the classification.
In both cases, it should be concluded that classification is strongly dependent on the EEG of the patient and could be a sign of overfitting of the model.

For completeness, the heatmaps relating to the last convolutional layer are also added to show how the spectrum is analysed in order to extract the features from the spectrogram images.

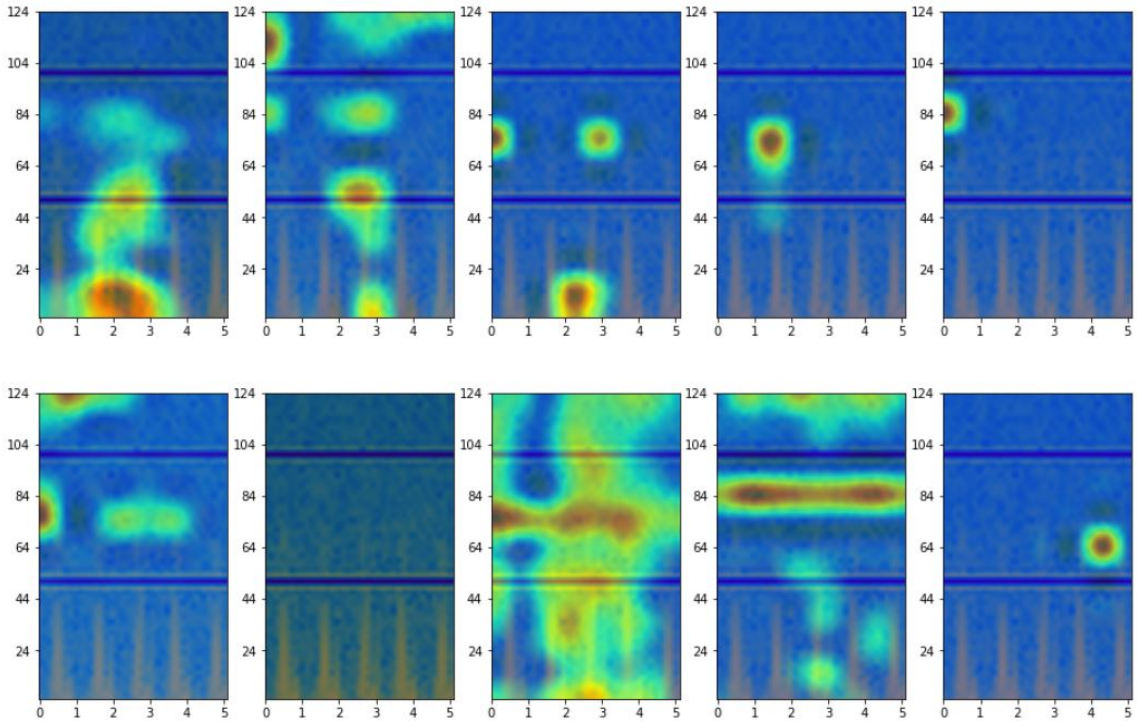


Figure 3.3: Heatmaps on spectrogram

Chapter 4

Conclusions and future works

Using the EEG of epileptic patients,an attempt was made to build a predictive model with the aim of identifying the temporal distance that separates the subject from a new epileptic seizure.
Although at first the results were encouraging (88 % accuracy in the random test) it must be concluded that despite the attempts made, it was not possible (at least for the time being) to achieve a learning threshold that goes far beyond pure chance (approximately 28 % in new subjects).

Perhaps, as highlighted by the last two CMs, the model is too trained on already seen patterns and is not able to *generalize* learn to new patients. In other words the model is affected by overfitting and a possible reason could be the few ES available for some patient (like Bal Urb) that imbalance too much the models weight. Moreover, some limitations of this research can be identified that probably affected the final results:

1. *On the starting datasets it's not known which seizures have been recorded.*

Without knowing the type of epileptic seizures of the patients, it is difficult to train a model that aims to predict new epileptic seizures: their origin is probably different from those on which the model is trained.
Segmenting each EEG into short time period like 5 s was an attempt to overcome this issue: the idea to get more data from each TRC file points to balance the lack of real ES patterns with more *cut* time series,hoping to extract some general valid features .
Of course a *real* solution could be the acquisition of a larger number of patients, which would reduce the risk of excessive 'specialization' of training only on certain types of seizure.

2. *The number and type of channels used during data recordings varies too much from patient to patient.*

As already mentioned in the preprocessing section, the fact that several patients had up to hundreds of channels and others far fewer did not facilitate the prediction task: having the same number and the same electrodes were maintained for all subjects, probably the extraction of features and therefore the learning would have been deeper.
Actually it's not enough to choose a *quite homogenous data* included all the TRC files with a n of channels belong into a short range like 24-35.

3. *EEG channels follow the same preprocessing as the no EEG ones.*

Both the EEG and no-EEG channels fall into the 24-35 range used during the analysis, but they actually followed the same pipeline as if they were of the same type. These happened for practical reasons because no specific channel list is available beforehand, and so it would be necessary to check, for each TRC, if each channel belongs to EEG class, ECG class, EMG class, etc. Moreover, using only the EEG channels should be difficult to gain much information because they are few. For sure, a detailed preprocessing of channel type could improve CNN performance, maybe adding a specific weight to one channel class than to the other.

4. *Code efficiency and more computational power.*

Having to manage a large amount of data ($\propto 10^4$ three-dimensional arrays of size equal to [62, 39, 35]) a better management of the remote pc resources, made available for the project, it would certainly have allowed a greater number of attempts and therefore better results could have been obtained. Of course, seeing that each training takes a lot of time, using more powerful machine or more computers could be another way to explore a more advanced deep learning architecture.

Bibliography

[1] Elizabeth A.Kelvin et al. "Prevalence of self-reported epilepsy in a multiracial and multiethnic community in New York City". In: *International Journal of Neural Systems* 77,Epilepsy Research (Dec. 2007), pp. 141–150. DOI: https://doi.org/10.1016/j.eplepsyres.2007.09.012.

[2] Saeid Abdi-Sargezeh Bahman Sanei. "Advances in epilepsy monitoring by detection and analysis of brain epileptiform discharges". In: *Psychology and Neuroscience* (2021). DOI: 10.1037/pne0000275.

[3] John F. Annegers, W. Allen Hauser, and Lila R. Elveback. "Remission of seizures and relapse in patients with epilepsy". In: *PubMed* Epilepsia (Dec. 1979), pp. 729–737. DOI: https://doi.org/10.1111/j.1528-1157.1979.tb04857.x.

[4] Haleema Anwar et al. "Epileptic seizures". In: *Discoveries (Craiova)* (May 2020). DOI: 10.15190/d.2020.7.

[5] Stéphane Auvin. "A common language of seizures and epilepsies: International League Against Epilepsy 2017 classifications". In: *PubMed* Dev Med Child Neurol (Mar. 2018). DOI: https://doi.org/10.1111/dmcn.13694.

[6] Ettore Beghi. "The Epidemiology of Epilepsy". In: *Review* Neuroepidemiology 54 (Mar. 2020), pp. 185–191. DOI: 10.1159/000503831.

[7] Ettore Beghi, Giorgia Giussani, and Josemir W Sander. "The natural history and prognosis of epilepsy". In: *PubMed* Epileptic Disorders (Mar. 2015), pp. 243–253. DOI: 10.1684/epd.2015.0751.

[8] Ettore Beghi and Dale Hesdorffer. "Prevalence of epilepsy—An unknown quantity". In: *PubMed* Epilepsia (Jan. 2014), pp. 963–967. DOI: 10.1111/epi.12579.

[9] A T Berg and S Shinnar. "The risk of seizure recurrence following a first unprovoked seizure: a quantitative review". In: *PubMed* Neurology (July 1991). DOI: 10.1212/wnl.41.7.965.

[10] A T Berg and S Shinnar. "The risk of seizure recurrence following a first unprovoked seizure: a quantitative review". In: *PubMed* Neurology (July 1991). DOI: 10.1212/wnl.41.7.965.

[11] Martin J Brodie et al. "The 2017 ILAE classification of seizure types and the epilepsies: what do people with epilepsy and their caregivers need to know?" In: *PubMed* Epileptic disord (Apr. 2018), pp. 77–87. DOI: 10.1684/epd.2018.0957.

[12] R Dhinakaran and Devendra Mishra. "ILAE Classification of Seizures and Epilepsies: An Update for the Pediatrician". In: *PubMed* Indian Pediatr (Jan. 2019), pp. 60–62.

[13] Resmi Cherian e E. Gracemary Kanaga. "Theoretical and methodological analysis of EEG based seizure detection and prediction: An exhaustive review". In: *Journal of Neuroscience Methods* (Jan. 2022). DOI: https://doi.org/10.1016/j.jneumeth.2022.109483.

[14] Jessica J Falco-Walter, Ingrid E Scheffer, and Robert S Fisher. "The new definition and classification of seizures and epilepsy". In: *PubMed* Epilepsy Res (Jan. 2018), pp. 73–79. DOI: 10.1016/j.eplepsyres.2017.11.015.

[15] Kirsten M Fiest et al. "Prevalence and incidence of epilepsy: A systematic review and meta-analysis of international studies". In: *PubMed* (Jan. 2017), pp. 296–303. DOI: 10.1212/WNL.0000000000003509..

[16] Robert S Fisher and Anna M Bonner. "The Revised Definition and Classification of Epilepsy for Neurodiagnostic Technologists". In: *PubMed* The Neurodiagnostic Journal (Mar. 2018), pp. 1–10. DOI: https://doi.org/10.1080/21646821.2018.1428455.

[17] Robert S. Fisher et al. "Operational classification of seizure types by the International League Against Epilepsy: Position Paper of the ILAE Commission for Classification and Terminology". In: *PubMed* Epilepsia,58 (Mar. 2017), pp. 522–530. DOI: https://doi.org/10.1111/dmcn.13694.

[18] Cynthia Harden et al. "Practice guideline summary: Sudden unexpected death in epilepsy incidence rates and risk factors: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology and the American Epilepsy Society". In: *PubMed* Neurology (Apr. 2017), pp. 1674–1680. DOI: 10.1212/WNL.0000000000003685.

[19] Ranjan Jana and Imon Mukherjee. "Deep learning based efficient epileptic seizure prediction with EEG channel optimization". In: *Biomedical Signal Processing and Control* (July 2021). DOI: https://doi.org/10.1016/j.bspc.2021.102767.

[20] Kitti Kaiboriboon et al. "Incidence and prevalence of treated epilepsy among poor health and low-income Americans". In: *PubMed* Neurology (May 2013). DOI: 10.1212/WNL.0b013e318293e1b4.

[21] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Communications of the ACM* (May 2017). DOI: https://doi.org/10.1145/3065386.

[22] Francis Levira et al. "Premature mortality of epilepsy in low- and middle-income countries: A systematic review from the Mortality Task Force of the International League Against Epilepsy". In: *PubMed* Epilepsia (Jan. 2017), pp. 6–16. DOI: https://doi.org/10.1111/epi.13603.

[23] Tobias Loddenkemper et al. "A proposal for a five-dimensional patient-oriented epilepsy classification". In: *PubMed* Epileptic Disord (Dec. 2005), pp. 308–316.

[24] H O Lüders, R Burgess, and S Noachtar. "Expanding the international classification of seizures to provide localization information". In: *PubMed* Neurology (Sept. 1993), pp. 1650–5. DOI: 10.1212/wnl.43.9.1650.

[25] Hans Lüders et al. "Classification of paroxysmal events and the four-dimensional epilepsy classification system". In: *PubMed* Epileptic Disord (Feb. 2019), pp. 1–29. DOI: 10.1684/epd.2019.1033.

[26] Caroline K Mbuba et al. "The epilepsy treatment gap in developing countries: a systematic review of the magnitude, causes, and intervention strategies". In: *PubMed* Epilepsia (Sept. 2008), pp. 1491–1503. DOI: 10.1111/j.1528-1167.2008.01693.x.

[27] Alison O'Shea et al. "DEEP LEARNING FOR EEG SEIZURE DETECTION IN PRETERM INFANTS". In: *International Journal of Neural Systems* (Aug. 2021). DOI: 10.1142/S0129065721500088.

[28] Alison OShea et al. “Deep Learning for EEG Seizure Detection in Preterm Infants”. In: *Int J Neural Syst* (May 2021). DOI: 10.1142/S0129065721500088.

[29] M Placencia et al. “Epileptic seizures in an Andean region of Ecuador. Incidence and prevalence and regional variation”. In: *PubMed* Brain (June 1992), pp. 771–782. DOI: 10.1093/brain/115.3.771.

[30] Robert S.Fisher et al. “Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)”. In: *PubMed* Epilepsia (Apr. 2005), pp. 470–472. DOI: 10.1111/j.0013-9580.2005.66104.x.

[31] Shah T. Sarmast, Abba Musa Abdullahi, and Nusrat Jahan. “Current Classification of Seizures and Epilepsies: Scope, Limitations and Recommendations for Future Action”. In: *PubMed* Epilepsia (Apr. 2014), pp. 963–967. DOI: 10.7759/cureus.10549.

[32] Ingrid E Scheffer et al. “ILAE Classification of the Epilepsies Position Paper of the ILAE Commission for Classification and Terminology”. In: *PubMed* Epilepsia (Mar. 2017), pp. 512–521. DOI: 10.1111/epi.13709.

[33] Afshin Shoeibi et al. “Epileptic EEG Classification by Using Time-Frequency Images for Deep Learning”. In: *International Journal of Environmental Research and Public Health* (Apr. 2021). DOI: DOI:10.1142/S012906572150026X.

[34] Afshin Shoeibi et al. “Epileptic Seizures Detection Using Deep Learning Techniques:A Review”. In: *International Journal of Environmental Research and Public Health* (Apr. 2021). DOI: doi.org/10.3390/ijerph18115780.

[35] Simon D. Shorvon. “The etiologic classification of epilepsy”. In: *PubMed* Epilepsia (Mar. 2011), pp. 1052–1057. DOI: https://doi.org/10.1111/j.1528-1167.2011.03041.x.

[36] Matti Sillanpää and Dieter Schmidt. “Natural history of treated childhood-onset epilepsy: prospective, long-term population-based study”. In: *PubMed* Brain (Mar. 2006), pp. 617–624. DOI: https://doi.org/10.1093/brain/awh726.

[37] Olafur Sveinsson et al. “The incidence of SUDEP: A nationwide population-based cohort study”. In: *PubMed* Neurology (June 2017). DOI: 10.1212/WNL.0000000000004094.

[38] Prasanth Thangavel et al. “Time–Frequency Decomposition of Scalp Electroencephalograms Improves Deep Learning-Based Epilepsy Diagnosis”. In: *PubMed* (Aug. 2021). DOI: 10.1142/S0129065721500325.

[39] David J Thurman et al. “Standards for epidemiologic studies and surveillance of epilepsy”. In: *PubMed* Epilepsia (Sept. 2011), pp. 2–26. DOI: https://doi.org/10.1111/j.1528-1167.2011.03121.x.

[40] David J Thurman et al. “The burden of premature mortality of epilepsy in high-income countries: A systematic review from the Mortality Task Force of the International League Against Epilepsy”. In: *PubMed* Epilepsia,58 (Nov. 2016), pp. 17–26. DOI: 10.1111/epi.13604.

[41] Yanli Yang et al. “Epileptic Seizure Prediction Based on Permutation Entropy”. In: *Frontiers in Computational Neuroscience* (July 2018). DOI: doi:10.3389/fncom.2018.00055.