Politecnico di Torino

Master Thesis in Biomedical Engineering

A.A. 2022/2023

Multi-class Motor Imagery classification technique based on a hybrid Deep Learning algorithm

Supervisor:

Luca Mesin

luca.mesin@polito.it

Candidate:

Rosario Spadaro

s267406@studenti.polito.it

This page was intentionally left blank

Summary

Abstract 6
1 Introduction 8
2 Nervous system 10
2.1 Neuron 12
2.2 Glial cells 15
3 Brain 18
3.1 Overview of the brain anatomy 18
4 EEG 27
4.1 Sources of EEG 27
4.2 EEG frequency bands 28
5 Brain Computer Interface (BCI) 31
5.1 BCI neural interface 31
5.1.1. Microelectrode array 32
5.1.2. Electrocorticography (ECoG) 34
5.1.3. Scalp EEG 35
5.2 Existing BCIs 37
5.2.1. P300 based BCI 37
5.2.2. Sensory Motor Rhythms (SMRs) based BCIs 38
5.2.3. Steady-State Visual Evoked Potentials (SSVEPs) based BCIs. 39
5.2.4. Slow Cortical Potentials (SCPs) based BCIs 40
6 Thesis objective 42
7 Materials and methods 42
7.1 Dataset 42
7.2 Montage 42

7.3 Experimental protocol 43
7.4 Proposed framework 47
7.5 Pre-processing 48
7.6 Epoching 50
7.7 Channel selection 51
7.8 Model architecture 53
7.8.1. Convolutional Neural Networks (CNNs) 54
7.8.2. Long-Short Term Memories (LSTMs) 57
7.8.3. Gate Recurrent Unit (GRU) 58
7.8.4. Hybrid Deep Learning (hDL) architecture 61
7.9 Evaluation performance 66
7.9.1. Leave N Subjects Out 66
7.9.2. Precision, Recall and F1 score 68
8 Results 69
9 Conclusions 71
10 Bibliography 73

This page was intentionally left blank

Abstract

Ischaemic or haemorrhagic strokes and injuries are the main cause of motor or cognitive disability. Patients become incapable to performing the most common daily activities. Treatment and therapeutic aids are an expensive solution. Rehabilitation therapies based on the growing use of BCIs have proven to be effective. BCIs of this type apply a repetitive mechanical or electrical stimulus to a impaired body part. The stimulus must be applied within a time range from the imagined movement of the dysfunctional limb. The aim is to restore a connection between the central nervous system and the area of the body that has lost its function. The application of a stimulus, mechanical or electrical, within a time window, facilitates motor recovery. Because the stimulus to be applied, a BCI must be able to recognise the kind of movement. The aim of this thesis is the classification of two imaginary movement using systems based on the most innovative deep learning techniques.

The reason for using deep learning architectures is to facilitate the recognition of imaginary motion by applying only a minimal pre-processing strategy. The dataset used is Physionet dataset, which contains EEG signal acquisitions from 109 subjects.

Afterwards, the EEG signal was only band-pass filtered and divided into the time epochs containing the imaginary movement. A hybrid Deep Learning network based on convolutional neural network and recurrent neural network was trained to extract the spatial and temporal features.

Finally, the evaluation of the network classification was performed by adopting the Leave-N-Subject Out method, where an evaluation is performed on N groups of subjects that were not used to train the network. Thus, due to the obtained performance, the proposed algorithm could be of interest for Brain Computer Interfaces applications.

1 Introduction

Stroke, caused by an ischemic or haemorrhagic brain injury, is the main cause of motor and functional impairment. The quality life of a person after a stroke changes negatively. The patient requires care and support in daily life. Moreover, treatment is not cheap either for families or for the health care system [1] [2].

Compensation or replacement of lost motor function can sometimes help stroke victims become more independent by modifying motor patterns or using aids. Starting therapy immediately after stroke to exploit neuroplastic mechanisms is the best strategy to maximise recovery [3]. New therapies have been proposed that seek to reactivate the brain's functional plasticity mechanisms and promote neuronal repair and regeneration in injured neural networks. One efficient and feasible way to stimulate the central and peripheral nervous system that could help reactivate functional plasticity mechanisms are BCIs. A BCI translates brain signals into input by providing a specific command as output. Thus, such a BCI can be used to stimulate paralysed limbs of the body by establishing a link between the brain and movement. These closed-loop neural interfaces activate neuroplastic mechanisms (e.g. Hebbian learning) [4]. The mechanism of neuroplasticity is relevant for stroke patients. BCI mediated motor recovery in stroke patients could therefore induce reconnection or reactivation between the central nervous system and peripheral pathways.

In the following chapters will be covered anatomical, physiological and functional principles that lead to the description of the process that

generates the electroencephalographic signal. In particular, the anatomy and functionality of microscopic structures, such as neurons and other small structures that make up the nervous system, will then be described, through to a description of the anatomy and functionality of macroscopic structures.

The detailed description of the microscopic structures of the brain is intended to help the reader understand the mechanisms that lead to the production of the electroencephalographic signal, while the understanding of the macroscopic structures serves to describe the areas of the brain from a functional point of view. Since the aim of the thesis is to study the classification of a multi class imagery voluntary movement, the knowledge of the cerebral cortex involved in the purpose will be further deepened. Then, based on knowledge of brain structures, the process leading to the generation of the EEG will be described. In addition, a description of the EEG acquisition instruments will follow. Finally, and not by order of importance, some types of Brain Computer Interface (BCI) will be described, deepening the knowledge of those operating through coupling with the motor cortex that is dedicated to the generation of voluntary movement.

2 Nervous system

A *neuron,* with the *glial cells*, is the main component of the nervous system. Neurons, through their excitability properties, are involved in the transmission of information in the nervous system and glial cells gives structural and metabolic support.

The nervous system is made up of the *Central Nervous System (CNS)* and the *Peripheral Nervous System (PNS)* (*Figure 2.1*).

CNS includes the *brain* and the *spinal cord*. It receives and process the information coming from the external environment, which are acquired by the sensory organs. Furthermore, receives and process also visceral information, that are information deriving from the internal organs. After processing the information, it sends instructions to the muscles and glands.

PNS consists of neurons that provide communication between CNS and the organs and can be subdivided into two divisions: afferent and efferent neurons.

Afferent neurons provide information from *somatic senses* (i.e. skin, muscles, joints), *special senses* (i.e. vision, hearing, equilibrium, smell and taste), and *visceral information* (i.e. blood pressure, blood pH).

Efferent neurons provide information from the CNS to the organs located in the periphery. Also, it can be divided into two categories depending on the muscles which innervate, that are: *somatic system* and *autonomic system.*

Somatic system regulates voluntary contractions of the skeletal muscles.

Autonomic system regulates involuntary organ contractions, such as sweat, glands or blood vessels and can be divided into two branches: *parasympathetic* and *sympathetic* nervous system [5].

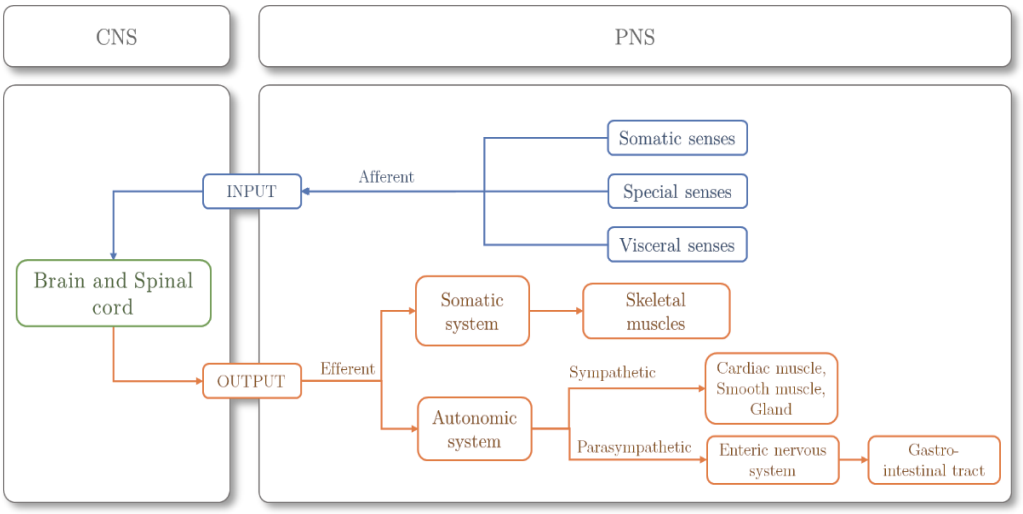


Figure 2.1 Schematic illustration of CNS and PNS. Blue lines represent afferent neurons and red lines represents efferent neurons.

2.1 Neuron

Neurons are made up of three main components: *cell body or soma*, *dendrites* and *axon (Figure 2.2).*

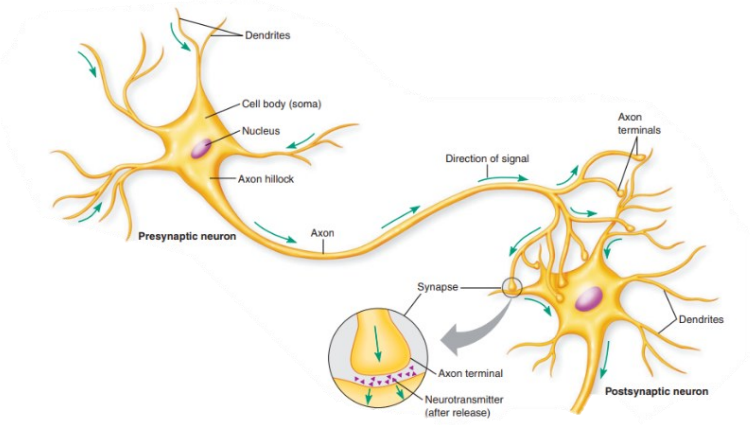


Figure 2.2. Example of a neuron anatomy and the coupling with another neuron. An illustration on how the signal flows from a pre-synaptic to the post-synaptic terminal neuron. The magnification represents a chemical synapse.

Into the *cell body* are performed the same activity performed in other cells, such as protein synthesis and cellular metabolism.

Dendrites are nervous fibers that branch out from a neuron. They receive information from other neurons at the synaptic junctions and transfer signals to the cell body. A *synaptic junction* is a gap between two neurons. It allows communication between nerve tissue cells.

Synapses are distinguished, according to their functionality, into:

- *Electrical synapses:* the signal is transmitted by direct passage of ions between two adjacent nerve cells,
- *Chemical synapses:* the signal is transmitted through the passage of neurotransmitter from the pre-synaptic terminal (axon

terminal that release neurotransmitter) to the post-synaptic terminal (dendrite, cell body of another neuron or cells of a receiving organ).

The *axon* is a branch of a neuron whose task is sending information to other neurons. Indeed, the signal flows from the cell body, propagates through the axon and reaches the terminal side of a neuron [6].

As mentioned above, the neuron is an excitable structure, and the information transmission is possible only through this property.

The membrane of the neuron, like that of other cells, is made up of a lipid bilayer.

Therefore, the excitability of neurons is given by the presence of channels onto the membrane. These constitute a passage for charged particles through it. Some channels are voltage-gated, that are trans-membrane protein that they pass from an open to a closed state and vice versa, due to a variation of the potential onto the membrane [7].

The *resting potential* of the cell is negative and is equal to -70 mV. From the chemical point of view, it means that, inside the cell, the concentration of potassium K^+ ions are greater, while outside there is a greater concentration of sodium Na^+ ions.

When a trigger event inside the cell body occurs, if a *threshold potential* of about -55 mV is reached, the depolarization of the cell takes place.

Therefore, an *action potential* originates from the *axon hillock*, which is a part of the axon that emerges from the soma. The axon hillock is characterized from the highest amount of voltage-dependent sodium channels.

All the action potentials have the same characteristic, that are 100 mV amplitude and 1,5 ms duration.

Figure 2.3 shows the step of a membrane potential change when an action potential flows along the axon. *Depolarization* causes the opening of the Na^+ ion channels, allowing Na^+ influx into the cell. The increased positive charge within the cell causes the K^+ channels to open. The efflux of K^+ ions help to restore the membrane potential difference to the resting state, causing *repolarization* and *hyperpolarization*.

Active Na^+/K^+ pumps promote the return of the membrane potential to the resting value. Consuming ATP, these pumps allow excess sodium to be taken out of the cell by letting potassium enter it: for every 3 Na^+ ions that come out, 2 K^+ ions enter. In this phase, the membrane potential difference returns to the resting state. A refractory period begins, lasting approximately 2 ms. During this period, the K+ channels are closed and the cell cannot depolarize.

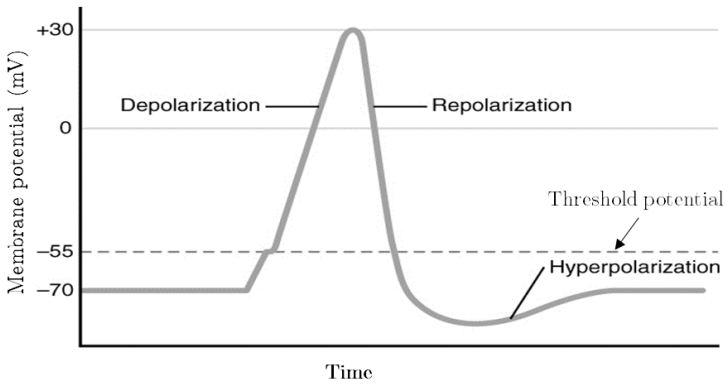


Figure 2.3. Illustration of a membrane potential trend with the flowing of an action potential

When an action potential arrives in the pre-synaptic terminal, it causes the release of chemicals, called neurotransmitters, contained in vesicles.

The neurotransmitter's effect on the post-synaptic neuron is determined by the neurotransmitter receptor it binds.

When the neurotransmitter coming from a presynaptic terminal binds to a postsynaptic terminal, a postsynaptic potential (PSP) is generated. The postsynaptic potential can cause the opening of channels that allow positive ions to enter, facilitating the generation of an Excitatory Post Synaptic Potential (EPSP). As described earlier, the cell body is innervated by multiple synaptic terminals. Pre-synaptic terminal releases a certain amount of neurotransmitter into the synaptic cleft. If neurotransmitters released into the synaptic cleft excites the nerve cell, EPSPs sum in space and time. If the threshold potential is reached, an action potential is generated. In a case which the released neurotransmitter binds to a receptors that cause the inhibition of an action potential, negative ions entry into the cell and a post-synaptic inhibitory potential (IPSP) is generated.

2.2 Glial cells

Glial cells constitute a large fraction of the brain components. Even if they don't produce electrical impulses, they still play an important role in complex cognitive functions. Indeed, glial cells have the role of giving a structure to neurons surrounding them, supplying nutrients and oxygen, insulating neurons from each other and removing toxic components for neurons.

Additionally, they play an important role in the preservation and consolidation of memories [8].

There are two groups of glial cells: *microglia and macroglia.*

Oligodendrocytes, Schwann cells, astrocytes, ependymal cells and radial glia are part of Macroglia.

Ependymal cells run though the spinal cord and ventricles of the brain. They are involved in creating CerebroSpinal Fluid (CSF).

Radial glial cells are progenitor, and the main function is to generate other cells, such as neurons, astrocytes and oligodendrocytes.

Astrocytes are the most abundant type of macroglia cells. They perform biochemical control of endothelial cells that form the blood–brain barrier (BBB), provide nutrients to the neurons, maintain extracellular ion balance and regulate cerebral blood flow [9].

Oligodendrocytes and *Schwann cells* form an insulation layer across axons, that is called *myelin sheath*. Oligodendrocytes form myelin sheath for CNS neurons and gives *white matter* its name. Instead, Schwann cells form myelin sheath for PNS neurons. Each Schwann cell provides myelin for only one axon, while an oligodendrocyte can myelinate many different axons [6].

Moreover, oligodendrocytes and Schwann cells are crucial to determine the velocity of electrical propagation of action potential in neurons. Indeed, they form an insulating layer of myelin sheath around the axons of the neurons *(Figure 2.4)*. Myelin sheath of nerves permits the action potentials to propagate faster and with a large amplitude with respect to neurons in muscle fibers that hasn't myelin sheath around the axon. The myelin sheath reduces leakage of ions across the cell membrane. However, gaps within the

myelin, called *nodes of Ranvier*, which contain voltage-gated sodium and potassium channels, permit the conduction along the entire axon. Because of this, the conduction takes a name of a *saltatory conduction*.

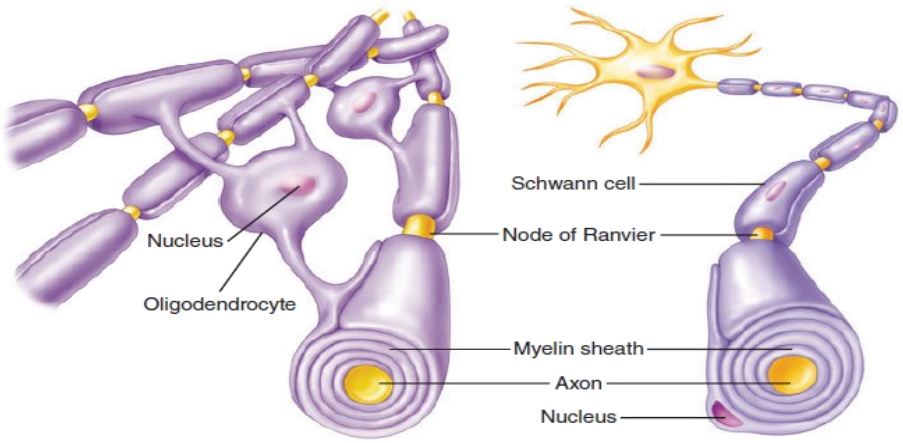


Figure 2.4. The figure to the left shows the myelin sheaths formed by oligodendrocytes in the central nervous system. Instead, the figure to the right shows the arrangement of myelin formed by Schwann cells in the peripheral nervous system. (Reproduced from [6]*).*

Microglia cells take care of the immune defense of the CNS. They are morphologically, immunophenotypically and functionally related to the monocyte/macrophage cells. They constantly search within the CNS for the presence of plaques, infectious agents and damaged neurons or synapses in order to remove them [10]. Furthermore, they play a key role in the search and elimination of pathogens which, by crossing the BBB, could cause inflammation of the nerve tissue.

3 Brain

3.1 Overview of the brain anatomy

In 1870 Eduard Hitzig and Gustav Fritsch applied an electrical stimulus to a region of a dog's brain. They saw a limb performed a movement. As a result, they found the relationship between the stimulated region and the area of the body dedicated to that function. Based on their experiments, Hitzig and Fritsch created a graphic representation of the primary motor cortex [11].

Electrical stimulation has also been used by Canadian neurosurgeon Wilder Penfield to map brain areas of patients undergoing neurosurgery [12]. The goal is to define areas of the cortex that have been damaged and have lost their motor or linguistic functionality. Furthermore, these studies led to the development of the *motor homunculus* of *Figure 3.1*, which is now present in all neurophysiology books.

Motor homunculus is a distorted map of the cortical area that generates the movement. More sensitive areas are represented with an exaggerate parts of the body, such as fingers and lower face. Less sensitive areas, which are shoulders and back, are represented with a tiny area on the cortex.

In addition to the primary motor cortex, Penfield identified the pre-motor cortex and the supplementary motor area.

In 1952, Woolsey and colleagues conducted animal experiments to map, not only the motor cortex, but also the sensory area of the cortex [13].

In the years of Woolsey's experiments, the first recording of a potential from a single neuron occurred on both anesthetized and awaked animals.

The experiment was carried out by inserting the tip of a microelectrode very close to the cortical neurons. Thus, it was possible to register the first action potential.

These experiments led to the division of the brain into different areas based on their functionality. Additionally, further histological and anatomical subdivision was possible.

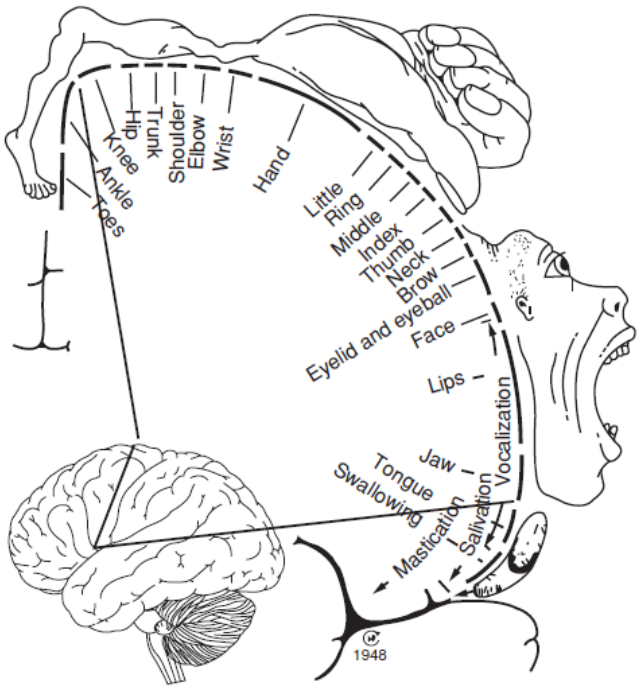


Figure 3.1. The motor homunculus derived by Wilder Penfield illustrating the effects of electrical stimulation of the cortex of human neurosurgical patients.

From the anatomical point of view, the brain is made up of two hemispheres coupled together by a large mass of nerve fibers called *corpus callosum*. Each hemisphere is covered by a *cortex*, which is a 1.5 to 4 mm thick structure and is known as *grey matter*. The grey colour is given by the presence of the cell body of neurons, dendrites and axons without myelin,

synapses and capillaries. The white matter, below the grey matter, is composed of nerves that connect the various areas of the cortex and the cortex with subcortical areas. Myelinated axons are responsible for the white colour of the matter *(Figure 3.2).*

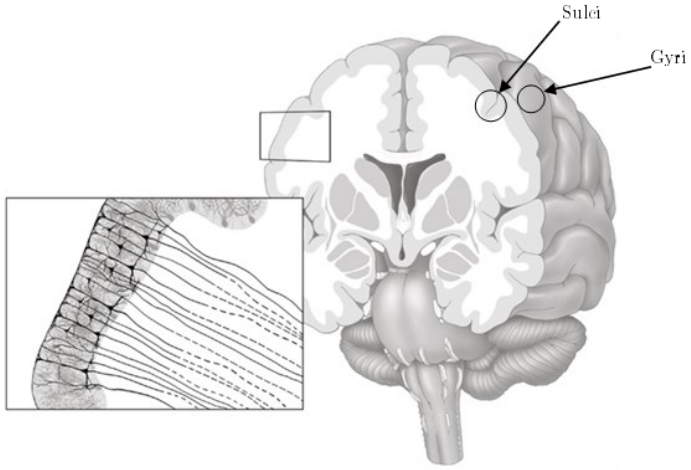


Figure 3.2. Illustration of the grey matter (in the outer side of the brain), the white matter (inner side of the brain), Sulci (grooves) and Gyri (ridges). (Reproduced from [16]*).*

Beneath the grey matter are other structures: *brainstem, cerebellum, diencephalon, thalamus and hypothalamus (Figure 3.3).*

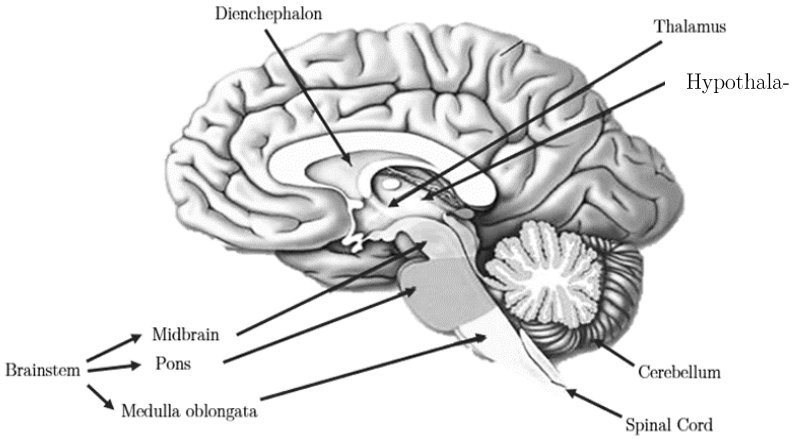


Figure 3.3. Anatomical view of four major component which constitute the brain. (Reproduced from [16]*).*

- *Brainstem* is a connection site between the spinal cord and the rest of the encephalon. Cranial nerves originate within it and branch out to the muscles, head, neck, chest and abdomen.
- *Cerebellum* is located at the back side of the brain. It is divided into two parts. The outermost one is made up of layers of folded tissue that contains neurons, while the innermost one contains nerve cells that communicate information from the brain. Cerebellum constitutes 10% of the total volume of the brain and contains about 80% of neurons of the brain [14]. It plays an important role in maintaining balance, in motor learning and in the coordination of voluntary movements by coordinating the strength and timing of activation of different muscle groups. In addition, it is also employed in some cognitive functions, such as language processing.
- The *diencephalon* is a component of the encephalon located between the brain and the brainstem and between the right and left hemispheres. It belongs to the SNC. The diencephalon comprises other structures, which are the thalamus and the hypothalamus.

  Its functionality is to establish a passage of information between the brain, brainstem and spinal cord. In addition, its function is related to learning, memory, and the regulation of the sleep-wake cycle.
- *Hypothalamus* is located below the thalamus. It regulates appetite, thirst, satiety, blood pressure, heart rate and sleep-wake rhythm.
- *Thalamus* is located between the brain and the brainstem and has an associative function between different cortical areas.

Morphologically, the cortex is made up of folds called *Gyri(ridges)* and *Sulci(grooves)* (*Figure 3.4*).

Sulci are present in the cerebral cortex of more evolved mammals and allow a greater volume of grey matter within the cerebral cortex without varying its thickness.

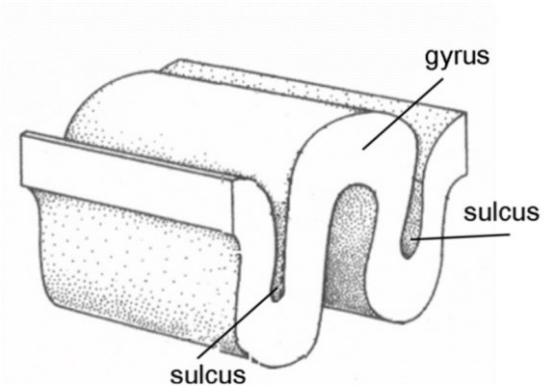


Figure 3.4. Illustration of Gyrus and Sulci

From a functional perspective, the cortex consists of four main components called *lobes: frontal, parietal, occipital and temporal*. Gyri and sulci separate the lobes.

Particularly, a *central sulcus* (CS) divides the frontal and parietal lobes. *Figure 3.5* is an illustration of how the cerebral cortex is divided into four areas based on their functionality.

Precentral gyrus and *postcentral gyrus* are part of, respectively, the anterior and posterior side of CS.

- *Frontal lobe* is the anterior portion of CS and extends into both hemispheres. It contains the *primary motor cortex (PMC), prefrontal association cortex, supplementary motor area and Broca's area*. Primary motor cortex lies along the anterior wall of CS and

continues into the precentral gyrus. Prefrontal area is involved in high-order executive function including complex cognitive behaviours, generation of voluntary movements, speech production and personality.

- *Parietal lobes* are positioned posterior to the central sulcus, superior to the temporal lobe and anterior to the occipital lobe. On the parietal lobes reside the *primary somatic sensory cortex*, which lies along the posterior wall of CS and continues into the postcentral gyrus, and *parietal-temporal-occipital association cortex*. Parietal lobes processing sensitive information coming from the skin.
- *Occipital lobe* is located behind and below the parietal lobe. It contains *primary visual cortex* and *secondary visual cortex*, which is highly specialized for processing information about static and moving objects, reading, written language comprehension. It is the area where visual stimuli are conveyed.
- *Temporal lobes* are mapped in the ventrally sides of the brain and border with the parietal lobes, occipital lobe and frontal lobes [15]. They are involved in higher-level visual processing, memory and auditory processing.

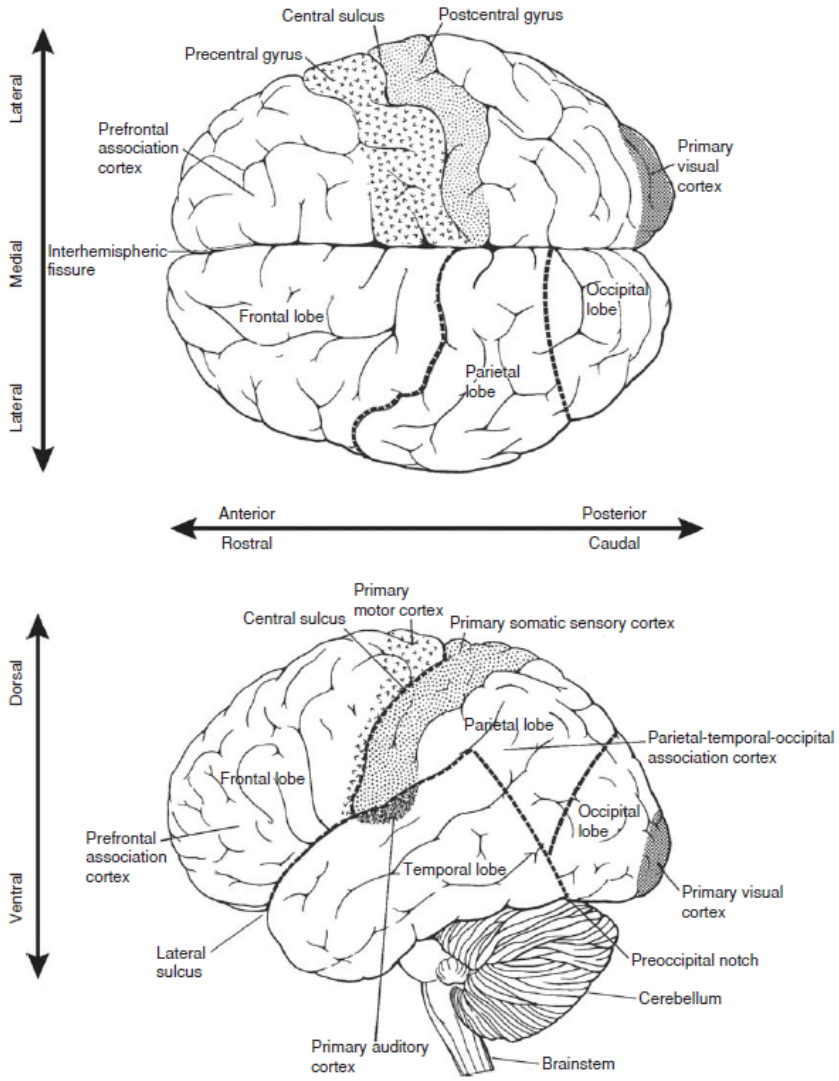


Figure 3.5. Dorsal (on top) and lateral views (on the bottom) of human cerebral cortex. (Reproduced from [16]*).*

According to a histological classification, the cerebral cortex can be further distinguished into three layers: *neocortex*, *paleocortex* and *archicortex*.

The *paleocortex* comprises a region at the bottom of the cerebrum that includes, but is not limited to, the olfactory cortex. The *archicortex* is a structure located deep within the temporal lobes. It plays a critical role in the formation of new memories and in spatial navigation. Lastly, the

neocortex represents the outer layer of the brain. It is involved in processes that include memory and learning. It is divided into six layers parallel to the cortical surface (*Figure 3.6*) [15]:

- Layer I is the *molecular* layer and contains very few neurons and is composed of the dendrites arising from pyramidal neurons in deeper layers and horizontally running axons.
- Layer II is the *external granular* layer and contains stellate cells and small pyramidal cells.
- Layer III is the *external pyramidal* layer and contains small and medium-sized pyramidal cells.
- Layer IV is the *internal granular* layer and contains many nonpyramidal neurons and receives much of the input coming from the cortex
- Layer V is the *internal pyramidal* layer and contains the largest pyramidal cells, the source of the long axons that project out of the cerebrum, which are named efferent fibers.
- Layer VI is the *multiform* layer and contains the greatest variety of cell types. It is the source of most efferent fibers which connect the cortex to the thalamus.

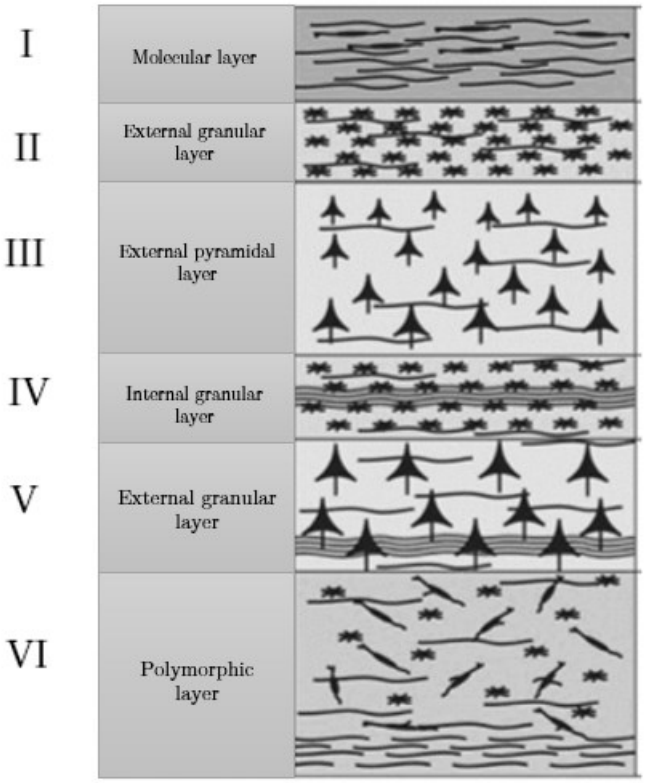


Figure 3.6. From top to bottom the figure illustrates from layer I to layer VI

4 EEG

4.1 Sources of EEG

The electroencephalographic (EEG) signal is the graphic representation in time of the potential difference between two different regions of the cerebral cortex.

The action potential and the Post-Synaptic Potential (PSP) are the two major contributions of the EEG signal, but most of the potential is constituted by the PSPs. This statement follows from the fact that most neurons that originate the action potential are not synchronized in time and the resulting signal amplitude is low. Then, the postsynaptic activity of pyramidal neurons is the true source of the EEG signal. In fact, although the amplitude of the PSPs is smaller, the time duration is longer (~30 ms) and the probability of synchronization of a large amount of neurons is higher [17].

The EEG detected on the scalp comes from cortical neurons.

The dendrites of cortical neurons arranged perpendicular to the surface of the cortex are crossed by a current that flow perpendicular to the surface.

Scalp EEG detects the current flow perpendicular to the cortical surface.

However, scalp EEG is a two-dimensional representation of a three-dimensional reality, which implies that it is not possible to determine the depth of the source [18].

4.2 EEG frequency bands

As aforementioned, EEG signals reflects the temporal synchronization of cortical pyramidal neurons. Due to its high spatial resolution, it is possible to extract a large amount of information from the frequency domain, including the presence of certain pathologies and disorders. the frequency spectrum of the EEG ranges from 0.1 Hz to approx. 100 Hz and can be subdivided into five non-overlapping frequency bands. The reason for the subdivision is the reduction of the time information. The frequency bands are:

- *Delta:* 0,1 - 4 Hz
- *Theta:* 4 - 8 Hz
- *Alpha (also called mu):* 8 - 13 Hz
- *Beta:* 13 - 30 Hz
- *Gamma:* $>$ 30 Hz

Each frequency band correspond to a mental state. The difference between the bands varies between subjects, age of the subjects and in the presence of pathologies. The phenomenon in which the areas of the cortex are synchronized is called Event Related Synchronization (ERS). In this case, the amplitude of the waves is high and the frequency is low. Conversely, if the frequency is high, cortical areas are not synchronized. The resulting amplitude value is low. The phenomenon of desynchronization of the activity of cortical areas is related to the initiation or suppression of an action and is called Event Related Desynchronization (ERD).

Delta band (0.1 - 4 Hz) is characteristic of infants and during sleep phases in the adult. If in adults it occurs when awake, it could be a case of a severe

brain disorder. In childhood, delta bands occur in the occipital lobe and, in adults, in the frontal lobe.

Theta (4 - 8 Hz) oscillations are often observed during the transition from wakefulness to sleep. An increase in theta power has been associated with the processing of emotional information and during memory-related tasks. In infants, theta frequency band is related to the expression of positive and negative emotions, feeding, drowsiness and modulation of attention [19].

Alpha band (8 - 13 Hz) is visualized in the occipital area. It is related to states of reflection and meditation and is most present during wakefulness.

It also occurs during eye closure.

Beta band (13 - 30 Hz) is present in the frontal region of the head. It is related to states of attention, concentration, alertness and anxiety.

Gamma band ($>$ 30 Hz) reflects the active maintenance of objects in memory. It is not normally present in the EEG tracing unless evoked by sensory stimulation.

A common scalp EEG shows frequencies until 40 Hz.

Only invasive neuronal interfaces can acquire an EEG trace with frequencies covering the entire frequency band, which extends up to 100 Hz.

EEG is essential because is an efficient method to diagnose neurological brain disorders such as epilepsy and autism spectrum disorder (ASD). For example, an EEG after a head injury, brain tumour or epileptic seizure presents slower waves than the waves of a normal EEG.

The Alzheimer's disease (AD) reduces the EEG complexity and changes the synchronization of the cortex. Furthermore, AD increases activity in

delta and theta bands and decreases activity in the alpha and beta bands [20].

The EEG in an epileptic disorder shows an abrupt rise time, a complex waveform and a slow-wave discharge that interrupts the continuity of the background rhythm [21].

5 Brain Computer Interface (BCI)

BCI systems constitute an interface between the user and the outside world. The user's brain signals are translated into an output. The user receives feedback on this output. The feedback affects the user's brain activity and influences the next output. Therefore, in a prosthetic arm control paradigm, the position of the arm after each movement will influence the person's intention for the next movement and subsequent brain signals with that intention [16].

5.1 BCI neural interface

The neural interface of BCIs is a hardware that detects the signal and translates it into useful information for the BCI system.

In order to be used with BCIs, they must fulfil certain characteristics. In fact, they must be reliable, safe and the degree of invasiveness must not be more than necessary.

Considering the above requirements, BCI neural interfaces can be divided into three categories:

- Electrodes placed on the scalp for EEG detection. They are non-invasive and record a low information potential from a large set of neurons and synapses,

- Electrocorticographic (ECoG) electrode arrays that are surgically implanted onto the cerebral cortex. They permit to record a moderate information from smaller sets of neuron and synapses,
- Microelectrode arrays surgically implanted into the cerebral cortex and record local field potential from a localized sets of neuron and synapses. They acquire a wide range of information.

5.1.1. Microelectrode array

Microelectrode array are devices containing several microelectrodes that allow neural signals to be obtained or transmitted to an external device.

The best-known microelectrode array is the Utah electrode array *(Figure 5.1)*. It consists, from the brain side, of many small electrodes which, from the way they are arranged, are called "bed of nails".

Invasive systems, such as the microelectrode array, are a good solution when it must acquire signals with a high information content for controlling prosthetic systems with high degrees of freedom, but they have some limitations. Regarding biological coupling, there are some problems arising from the use of invasive electrodes and the surrounding biological tissue.

Tissue response to implantation depends on the size of the microelectrode, the distance between the stems, the composition of the material and the time of insertion. Tissue response is typically divided into short-term response and long-term response. The short-term response begins with an increase in the population of astrocytes and glial cells surrounding the device.

The recruited microglia initiate inflammation and begin a process of phagocytosis of the foreign material. Over time, the astrocytes and microglia recruited to the device begin to accumulate, forming a sheath surrounding the microelectrode.

The sheath insulates the electrodes and increases impedance.

To minimize the negative effects of device insertion, the surface of the device can be coated with proteins that promote neuron attachment, such as laminin, or substances that elude drugs [22].

Implantable devices that work with invasive electrodes include deep brain stimulators (DBS).

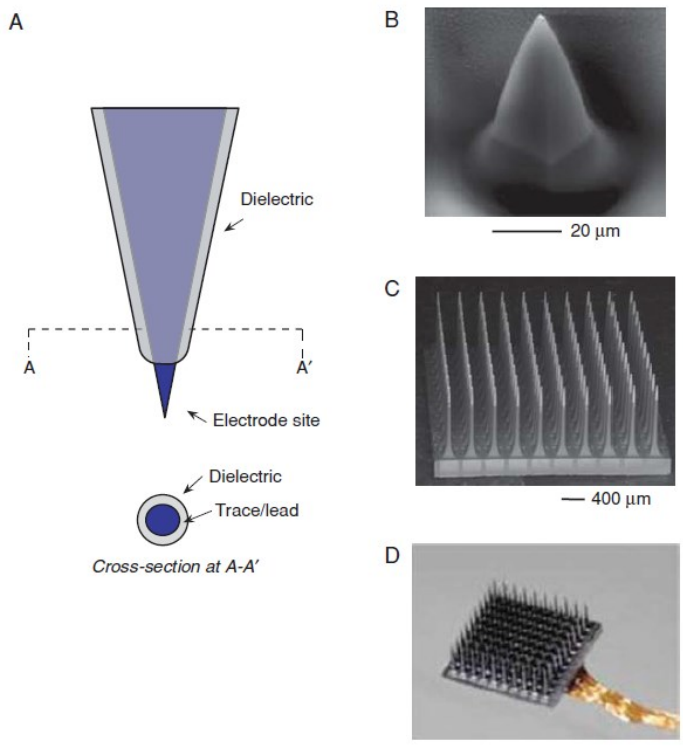


Figure 5.1. (A) Illustration of the structural components near the distal end of a single tooth of the Utah microelectrode. The uninsulated tip is the electrode site. The insulated part of the tooth is the dielectric component. (B) electron scanning myograph of a tooth. (C) Scanning electron micrograph of the bottom side of the Utah microelectrode. (D) Utah microelectrode connected to a gold wire bundle. (Reproduced from [16])

DBS are effective in treating the movement disorders of Parkinson's disease, cochlear implants which work with auditory nerve stimulation and cardiac pacemakers [23]. The use of MEAs could help restore vision by stimulating the optic nerve [24].

5.1.2. Electrocorticography (ECoG)

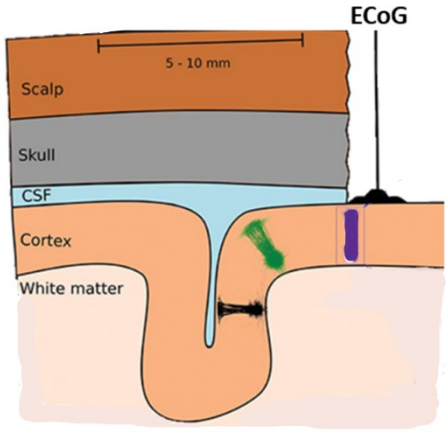


Figure 5.2. Location of implantation of a ECoG electrode for the detection of a local field potential.

ECoG electrodes consist of a series of electrode plates placed on a thin silicone sheet on the surface of the brain directly under the dura or placed inside a groove *(Figure 5.2)* [25]. They are capable of acquire a local field potential with a high SNR ratio.

The higher SNR is due to the shielding effect of the skull reducing ambient and muscle noise and the proximity to the source of the potential.

Biological response to the implant is reduced due to location of the electrodes.

The ECoG was applied for epilepsy surgery. Also, it is possible to use the ECoG as a monitoring system for seizures and for controlling simple cursor movements [26].

5.1.3. Scalp EEG

Scalp EEG is a type of non-invasive EEG acquired from electrode placed on the scalp.

EEG signal is acquired from the scalp using electrodes of varying numbers.

A multi-channel montage system consists of a minimum of 32 electrodes mounted on a cap. Each electrode is placed at a fixed distance from the surrounding electrodes. The distance is defined by an international standard.

For example, the most commonly used standard is 10-20 international standard where "10" and "20" are proportional distances of respectively 10% and 20% of the total length along contours between skull landmarks. "20" refers to the distance between an electrode and the adjacent electrodes and "10" refers to the distance between nasion and the first adjacent electrode and between the inion and the first adjacent electrode *(Figure 5.3). Nasion* is the point between the eyes and *Inion* is a point of back of the skull.

Electrode placements are labelled according to the brain areas: F (frontal), pre-frontal (Fp), C (central), T (temporal), P (posterior), O (occipital) and Z sites (Fpz, Fz, Cz, Oz). Z letter denotes the reference point. A letter refers to the bone prominent behind the ears and, sometimes are used for contra-lateral referencing. T3-T4, P3-P4 are used for seizure detection montage.

The letters are followed by odd numbers at the left side of the head and with even numbers on the right side.

Each electrode is an Ag-AgCl disk of 1 mm to 3 mm in diameter, with long flexible leads that can be plugged into an amplifier [27].

It is possible to acquire EEG signal from the scalp with higher resolution

systems. In this case, the number of electrodes increase fulfilling the empty space of the international standard system, becoming 10-10 system or 10-5 system. The amplitude of EEG signal recorded from the scalp is about 10 µV to 100 µV.

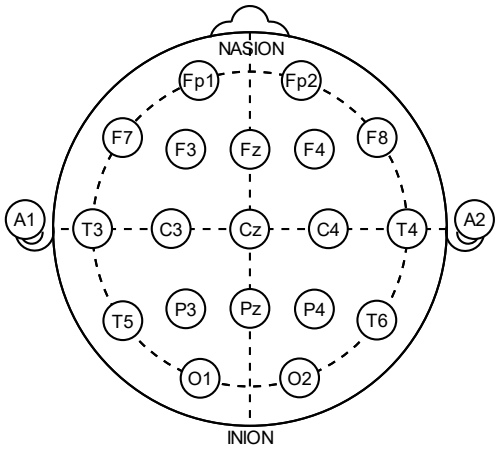


Figure 5.3. International Standard 10-20 system

EEG deriving from the scalp has low spatial resolution and high temporal resolution but is very useful in many applications, such as diagnoses or treatment of the following disease:

- Brain tumour
- Brain damage from head injury
- Encephalopathy
- Stroke

- Sleep disorders

However, the EEG is not only used for clinical applications. As its use has become more widespread, its fields of application have widened to including cognitive neuroscience, psychiatric disease diagnosis, neuromarketing, neurorobotics, sports science and human brain mapping. Recently, scalp EEG can be used for real-time applications such as BCI and neurofeedback [17].

5.2 Existing BCIs

It is common to see BCIs used in certain television scenes, where they are depicted as an enhancement of human capabilities or, as a portal to immersion within a virtual reality, as in *The Matrix* or *Ready Player One*. In reality, BCI has a wider field of application in the rehabilitation or support of people with disabilities or physical impairments.

5.2.1. P300 based BCI

ERPs are a manifestation of neural activity in the brain. They occur in relation to an internal or external triggering event. Reaction to the stimulus can be exogenous or endogenous. Exogenous reactions have a short latency period and usually occur no more than 150 ms after the triggering event.

Their localization on the scalp, latency and morphological characteristics depends on the type of stimulus applied.

Endogenous responses are not stimulus-dependent but are due to cognitive activities. P300 is an endogenous response. It occurs between 250 and 750

ms after the triggering event [28] [29] [30]. The variable latency is since the P300 ERP is modulated by an unconscious decision process and by the difficulty of the stimulus [31]

It is originated in relation to an Oddball paradigm. Oddball paradigm provides two classes of stimuli (visual or auditory), which are delivered within a sequence in a non-repetitive manner. One of the two stimuli is less frequent. The less frequent stimulus represents the Oddball stimuli.

P300 occurs in the signal taken from the scalp as a positive trend followed by a negative trend. It is maximal on the Pz electrode and extends over the entire posterior parietal region. BCIs based on P300 stimuli are non-invasive and can be reparametrized within a few times.

P300 based BCIs applications are manifold. They are used to provide a communication system for ALS patients, in the context of virtual reality or, even in gaming application. For example, P300 based BCIs can move a cursor in one of four directions instead of selecting item [32], enable control of a mobile robot or to control a wheelchair [33] [34].

5.2.2. Sensory Motor Rhythms (SMRs) based BCIs

SMR are electrical or magnetic oscillations recorded on the sensorimotor cortex. SMR, in the scalp EEG, falls in the mu rhythms (8-12 Hz) and beta rhythms (18-30 Hz). In the ECoG and MEG recording it is possible to find, respectively, electrical and magnetic activity of the sensorimotor cortex also in the gamma frequencies (30-200 Hz). In this case, an imaginary or executed movement is the event that determines the response. Many studies have

shown that executed or imagined movement causes a desynchronization in the mu/beta frequency (ERD) and a synchronization in the mu frequency (ERS).

This means that the power of the mu/beta rhythms decreases during the imagining or execution of the movement, while the power of the mu rhythms increases after the execution or imagination of the movement [35].

BCI based on SMR signals have been used in various healthcare applications, such as neurological rehabilitation of a stroke brain area.

BCI systems coupled to robotic assistive devices have shown promising results for stroke rehabilitation, however to date none of these systems are used in clinical practice [36]

BCI working with this type of signal has led to promising results in motor rehabilitation interventions and can be applied to patients without residual movement [37].

Furthermore, they can be used for cursor control and robotic wheelchairs for people who have lost the ability to walk.

5.2.3. Steady-State Visual Evoked Potentials (SSVEPs) based BCIs

Visual Evoked Potentials is a potential elicited by a visual stimulus, which is a flashing light or the appearance of an image. The components of the VEP are the N70 and P100, which are a negative deflection after 70 ms and a positive deflection after 100 ms from the stimulus, respectively [38].

Steady-State Visual Evoked Potentials (SSVEPs) are steady-state oscillations also evoked by a visual stimulus. Successive stimuli evoke the same response, and the superposition of responses produces a steady-state oscillation.

The paradigm involves displaying different boxes flashing at a different frequency on the screen. Each box is associated with a BCI output. The user only selects by looking at the box that performs a specific function. The potential is then read in the occipital region. Frequency of the stimulus is typically equal to the frequency at which the looking box flashes. SSVEP based BCI can control a functional electrical stimulator (FES) for knee flexion or the roll position of an aircraft in a flight simulator [39].

With this type of BCI it is possible to drive a car within a map in a virtual game environment [40]. In addition, it is possible to control an avatar in virtual reality. The avatar's movements depend on the box in which it is looking [41].

5.2.4. Slow Cortical Potentials (SCPs) based BCIs

SCPs are slow variations of the potential. They are recorded, such as SMR, over the sensorimotor cortex. SCPs are potentials occurring in relation to an event. They are linked to the time and phase of particular sensorimotor events (i.e. they occur at predictable times before, during or after certain events). Typically, it is a negative shift of the potential that precedes an imagination or planning of movement or a cognitive task (i.e. arithmetic calculations). SCPs are followed by a wave called *movement related*

potential. SCPs are Bereitschaftspotential (BP) and Contingent Negative Variation (CNV).

BP is a negative SCP that occurs 500-1000 ms before a self-paced movement.

CNV is a negative SCP that occurs 200-500 ms after a stimulus [42] [43] [44].

Similar to SMRs, SCPs and related potentials over sensorimotor areas are associated with motor imagination and real movements [45]. The paradigm used is based on two conditions. The first is a resting condition and the second is activity condition.

SCP-based BCIs are mainly used for cognitive tasks. Users using BCIs based on SCPs are submitted to long training periods consisting of repetitive sessions to enable them to control SCPs [46] [47].

An example of SCP-based BCIs is controlling the direction of a ball on a display by acquiring the signal from the Cz, C3 and C4 electrodes. Another example is the spelling system for people with disabilities, such as ALS persons [42].

6 Thesis objective

The aim of the thesis is to classify two classes of imaginary movements (MI) using an algorithm based on deep learning (DL). In addition, this algorithm must be able to provide a good level of performance even with minimal pre-processing of the raw data.

7 Materials and methods

7.1 Dataset

The dataset used in the study was made public by PhysioNet and the download was made possible through the official website (EEG Motor Movement/Imagery Dataset v1.0.0 (physionet.org)) [48].

7.2 Montage

The EEG signal is acquired through a helmet equipped with 64 electrodes at a sampling frequency of 160 Hz. The electrodes, as illustrated in *Figure 7.1*, are positioned according to the international 10-10 standard. Electrodes (Nz, F9, F10, FT9, FT10, A1, A2, TP9, TP10, P9 and P10) are excluded.

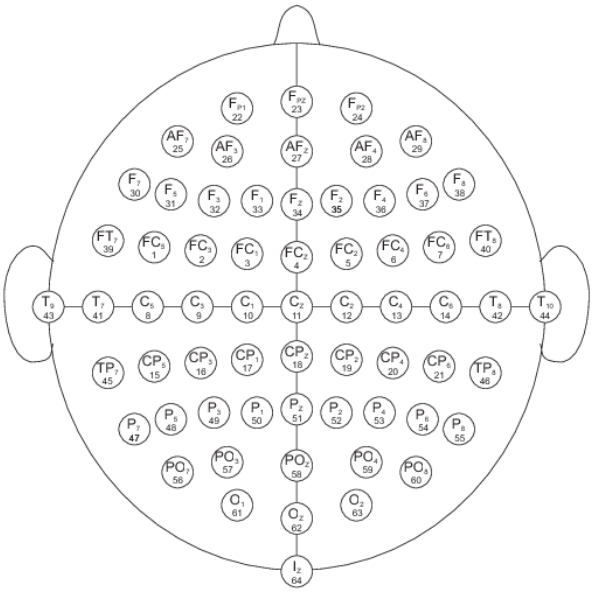


Figure 7.1: Illustration of the electrode montage according to the international 10 -10 system

7.3 Experimental protocol

The dataset under consideration consists of 109 subjects, each running 14 runs. The advantage of using a dataset consisting of numerous subjects is related to a better operation with DL networks. In fact, the numerosity of the observations avoids the problems of overfitting the network, saving on data augmentation operations. In addition, the large number of subjects makes it possible to obtain a network with good generalization capabilities.

In the first two runs, an EEG signal lasting one minute each was acquired. In the first run, the subject is asked to keep his/her eyes open while, in the second run, the subject keeps his/her eyes closed. In the remaining runs, lasting approximately 2 minutes, the subject is asked to perform or imagine

a movement, according to the following protocol consisting of the following 4 tasks:

1. A target appears, for 4.1 s, on the right or left side of the screen. The subject opens and closes the corresponding fist until the target disappears. Afterward, the subject relaxes.
2. A target appears, for 4.1 s, on the left or right side of the screen. The subject imagines opening and closing the corresponding fist until the target disappears. Subsequently, the subject relaxes.
3. A target appears, for 4.1 s, on the upper or lower side of the screen. The subject opens and closes, if the target is on the upper side, both fists and, if the target is on the lower side, both feet, until the target disappears. The subject then relaxes.
4. A target appears, for 4.1 s, on the upper or lower side of the screen. The subject imagines opening and closing, if the target is on the upper side of the screen, both fists. If the target is on the lower side of the screen, the subject imagines opening and closing both feet until the target disappears. Then the subject relaxes.

In this thesis, Movement Imagery (MI) is the object of the study. For each subject, only those runs related to MI are considered and the runs related to the execution of the movement are excluded. The runs studied are:

- Run 4 (R4)
- Run 6 (R6)
- Run 8 (R8)
- Run 10 (R10)

- Run 12 (R12)
- Run 14 (R14)

Each run, excluding the two first runs, includes the annotations T0, T1 and T2, which identify respectively resting state, beginning of movement (real or imagined) of the left fist (LF) or both fists (LR) and beginning of movement (real or imagined) of the right fist (RF) or both feet (BF).
For each run of the respective subject, observation windows of 4.1 seconds EEG signal length were extracted. It is possible to distinguish each observation as follows:

- If the T1 annotation is contained in R04, R08 and R12, the MI is classified as Left Fist (LF)
- If T1 annotation is contained in R06, R10 and R14, the MI is classified as Both Fists (LR)
- If T2 annotation is contained in R04, R08 and R12, the MI is classified as Right Fist (RF)
- If T2 annotation is contained in R06, R10 and R14, the MI is classified as Both Feet (BF)

Figure 7.2 shows a schematic subdivision of the four MI mentioned above.

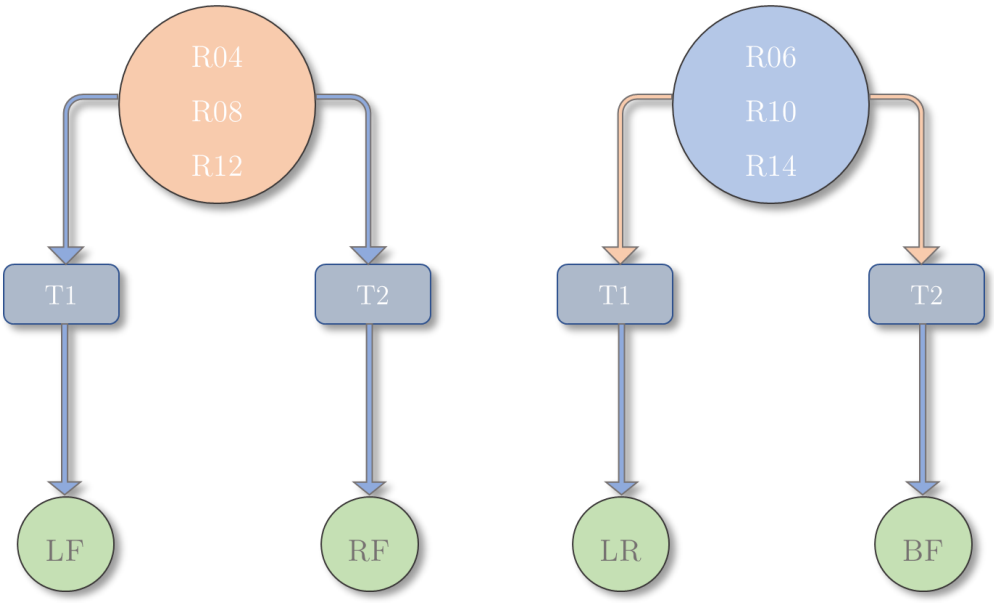


Figure 7.2. Subdivision of the observations according to the annotations and to the run.

Subjects S88, S92 and S100 were excluded from the study because they had a sampling frequency or observation window duration that did not agree with the dataset description.

Finally, only R04, R08 and R12 were considered for the study.

7.4 Proposed framework

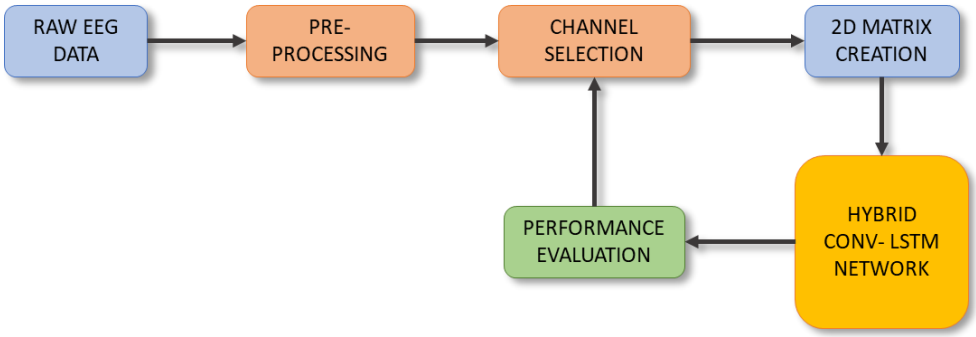


Figure 7.3. Workflow of the thesis process

The files containing the EEG signal were provided in the standard '.edf' format. The programming and numerical calculation platform called MATLAB® [49] was used to deploy the code. The first process is converting the files from the '.edf' format to a format easily manageable by the platform, i.e. '.mat' format. After conversion, a filtering task was performed in which, low frequency trends and information carrying the frequency band above 30 Hz were removed. For each subject, an epoching process was performed.

Epoching consists of the extraction of time samples containing the signal corresponding to the MI-EEG. Subsequently, in accordance with the subdivision in *Figure 7.2*, LF and RF labels were assigned. A 2D matrix of size $N \times 657$ was created, where the rows contain the number of channels and the columns the time points. In this case, N are the channels and 657 are the time samples. Hence, the bidimensional matrix constitutes the inputs of the neural networks.

In addition, a study was conducted in order to evaluate the performance of the classification in relation to the number of electrodes considered. *Figure 7.3* describes, in brief, the process adopted to obtain the results.

7.5 Pre-processing

Pre-processing step is a very important activity. It allows the selection of frequencies of interest excluding unnecessary ones. In addition, this process removes certain artefacts. For example, eye movement and eye - blinking generate electrical activity that constitutes an artefact in the EEG pathway. Most of the power of ocular artefacts is in the delta and theta band and decreases with increasing frequency, such that it is considered negligible in the alpha and beta band. Ocular artefacts are greatest in the region close to the eye and, as they move away from the eye, they decrease their intensity [50].

Another artefact is generated by electrical activity in response to facial muscle contraction. Typically, the frequency band of EMG artefacts, like muscle facial contraction, overlaps the frequency band of the EEG, but most of the power is above 30 Hz [51].

Electrical activity due to the QRS complex of the ECG also appears within the EEG trace.

The change in impedance or potential between the scalp and the electrodes, also causes artefacts that appear as slow oscillations around a frequency of 1 Hz.

Regarding the considerations addressed in Chapter 4 on ERD/ERS and the above-mentioned artefacts, the frequency bands involved in MI-EEG were isolated. Thus, the frequency band considered are those corresponding to the mu (8-12 Hz) and beta (13-30 Hz) rhythms. However, it has been observed that the addition of the delta and theta frequency bands leads to better results. Bandpass filtering was performed by implementing two Butterworth filters. The first filter performs high pass filtering with a cut-off frequency of 0.8 Hz. The second filter performs low pass filtering with a cut-off frequency of 30 Hz.

The Butterworth filter is an IIR filter and its use was favoured because it has a flat frequency response curve [52].

In order to better observe the power frequency distribution, the power spectral density (PSD) was calculated for Subject 6. *Figure 7.4* shows the PSD of the signal before filtering.

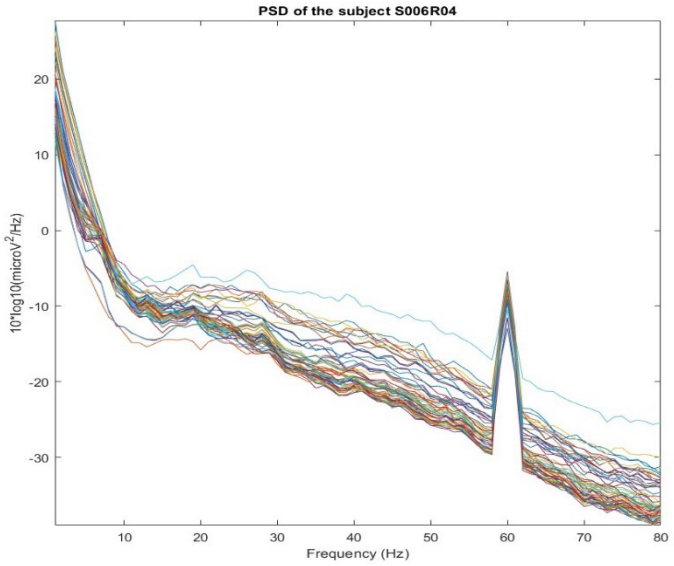


Figure 7.4. PSD illustration of Subject 6 before filtering

A power peak centred at 60 Hz can be seen. The peak corresponds to the power line interference. Moreover, it can be seen that the power is more distributed at low frequencies and decreases as you move away from the low frequencies.

Figure 7.5 shows the PSD of the signal after filtering. The signal power, after 30 Hz, i.e. after the upper cut-off frequency, decreases steeply.

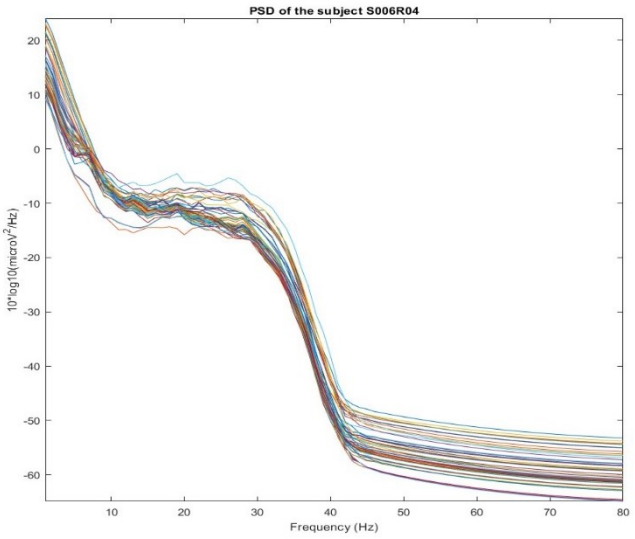


Figure 7.5. PSD illustration of Subject 6 after Butterworth filter (0,8-30 Hz)

7.6 Epoching

The segmentation of the EEG signal involves the extraction of time windows corresponding to the duration period of the cue. A time window of 4.1 s, in which the presence of the MI was expected, was considered.

The extracted time windows constitute the inputs to the neural network.

Thus, from the epoching, matrices of dimension $[N \times 657]$ were obtained. N corresponds to the number of channels and 657 are the sampling point. A 2D matrix previously created, is also called an observation.

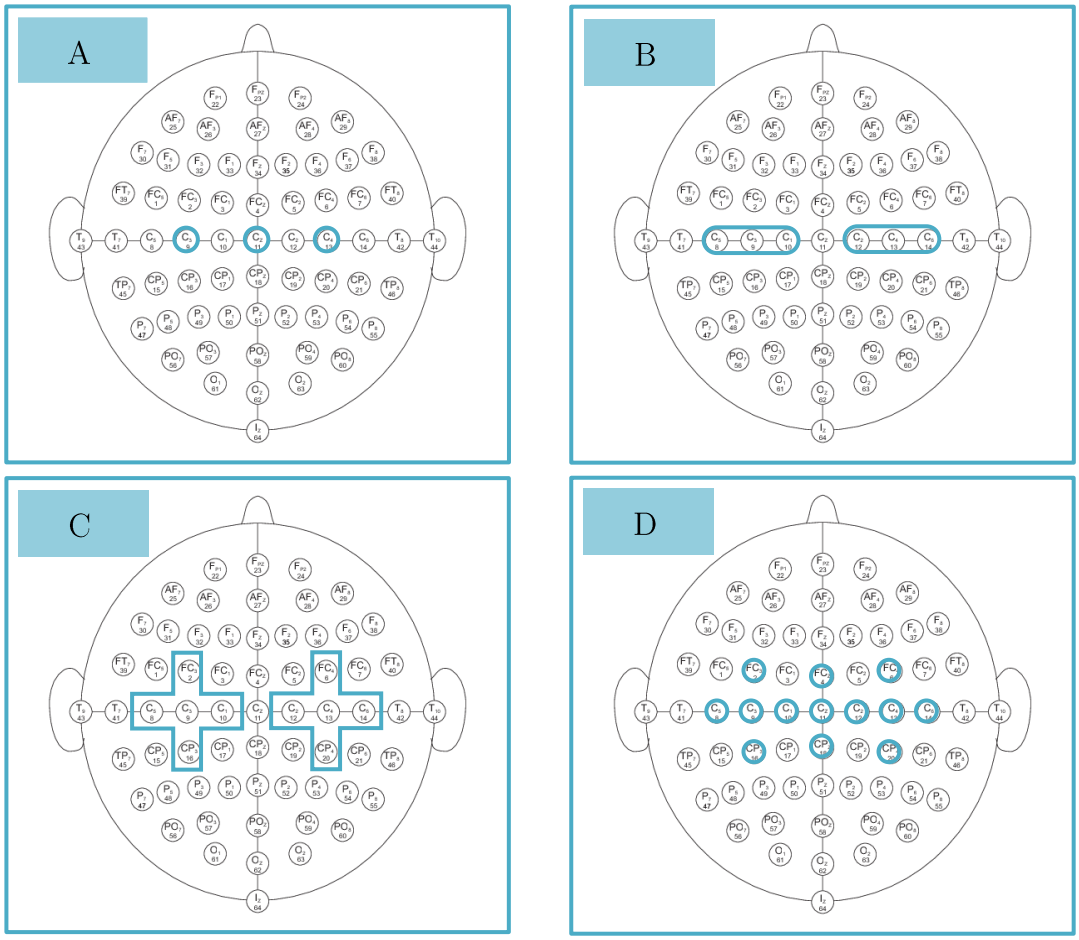
Each observation has an associated label. According to the dataset description, the labels are LF or RF and the total number of observations is 4748. LF was assigned for the MI-EEG of the left hand. RF was assigned for the MI-EEG of the rigth hand imagery movement.

7.7 Channel selection

The channels selected for the study are those present on the SMR cortex. Channels on the SMR cortex are less corrupted by ocular artefacts. In addition, considering only a few channels brings advantages to the positioning of the EEG helmet, leading to a reduction of the electrode placement time on the scalp. From the point of view of network training, a few numbers of electrodes decrease the training time of the neural network. However, it may lead the network to overfitting problems. As an opposite effect, the use of the entire number of channels leads to an increase in training time with a consequent increase in computational costs.

Some studies shows that different types of motor imagery have the most impact on the signals of C3, Cz, and C4 channels. Therefore, C3, Cz, and C4 channels was taken as the core and gradually expand the number of channels, reaching seven groups of channels.

The detail of the groups chosen are illustrated in *Figure 7.6*, which is a topographical representation of the channel groups on the scalp.


A
B
C
D

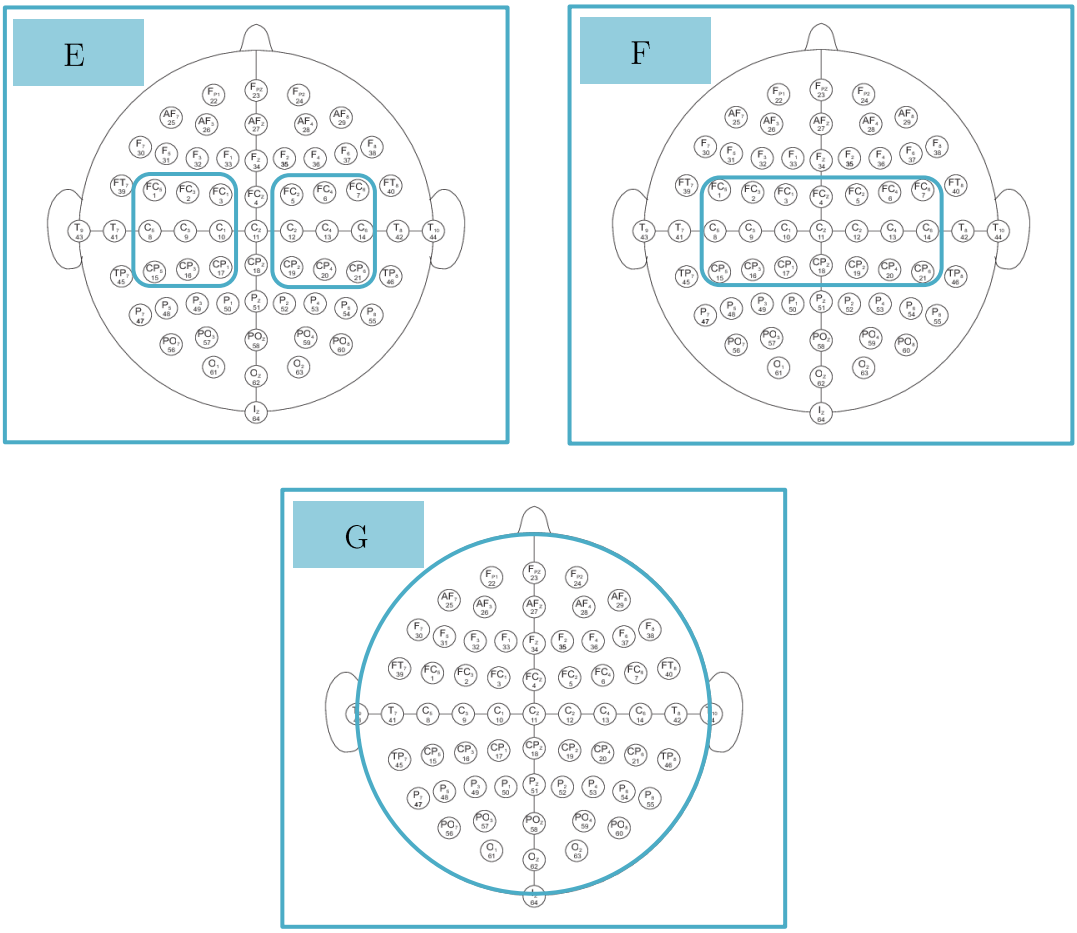


Figure 7.6. Topological distribution of channel groups on the scalp

7.8 Model architecture

For the analysis of the MI-EEG signals, alternative approaches to Machine Learning (ML) have been studied in recent years and are all based on the use of Deep Learning.

In particular, its ability to learn features from raw signals has been investigated.

Architectures as Convolutional Neural Networks (CNNs) have shown great success in the field of image and audio signal recognition [53].

The EEG signal, in the time domain, has a high variability between trials of the same subject (intra-subject variability) and between trials of different subjects (inter-subject variability) [54]. These features make the networks poorly efficient. Some studies, to solve this problem and obtain better classification results, extract features in the frequency domain and in the spatial domain and use them as input to the neural network.

One of the main objectives is to find formulation of inputs or neural networks that make the features extraction process as automatic as possible.

Deep Learning networks, such as the most famous Convolutional Neural Networks (CNNs), can do this.

In the following, a network that includes convolution layers will be called CNN and network that includes recurrent layers is called RNN.

7.8.1. Convolutional Neural Networks (CNNs)

CNNs have been applied in several fields, which include computer vision, speech and face recognition [55] [56].

The structure of CNNs was inspired from the structure of the human brain. It's based on the visual cortex [57].

A convolutional network takes its name from the convolutional layers.

A convolution layer consists of numerous neurons, which connect the input images or the images coming from the output of other layers. Thus, the layer learns the local characteristics of the image.

For each region, the dot product of the *weights* with the input is performed. A *bias* term is then added. A set of weights applied to the image region is called a *filter*. A filter is also called *kernel*. A filter repeats the same procedure for each region, i.e., it performs a convolution of the input (*Figure 7.7*).

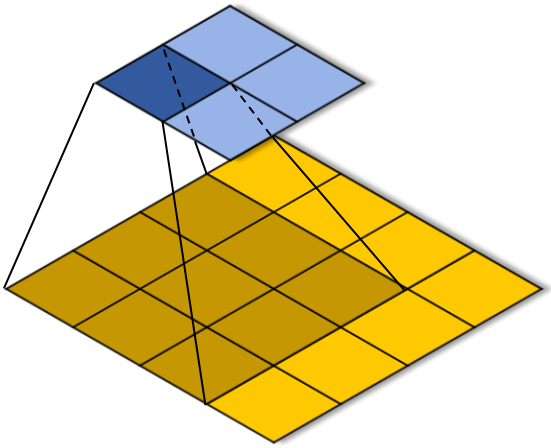


Figure 7.7. Illustration of the kernel sliding over an input and executing the scalar product.

The step size at which a filter moves on the image is called *stride*. For example, a stride of value 2 is represented as follows in the *Figure 7.8*.

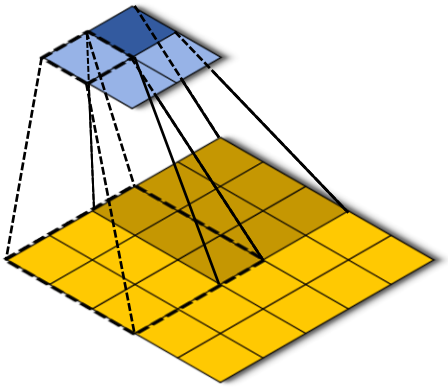


Figure 7.8. illustration of a filter with stride of value 2 sliding along an input

The number of weights in a filter is $[h \times w \times c]$, where h is the height, w is the width and c is the number of channels in the filter. If the input is a RGB image, c is equal to 3. In addition, the number of filters determines the output channels of a convolution layer. A filter moving over the input generates a *feature map*. The number of feature maps is equal to the number of filters, and each filter has a different set of weights and biases. Hence, the total number of parameters is $[(h \times w \times c + 1) \times Number\ of\ Filters]$, where 1 is the bias.

When the filter exceeds the size of the input, samples can be added to the edge of the input *(Figure 7.9)*. This procedure is called *padding* [58].

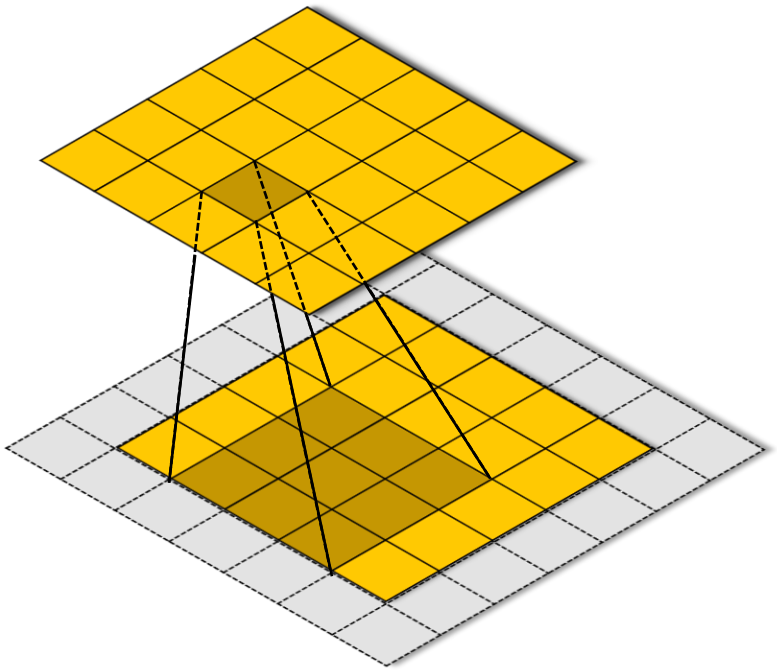


Figure 7.9. The grey rows and columns represent padding

7.8.2. Long-Short Term Memories (LSTMs)

An example of a RNN is the LSTM network, which is able to learn long-term dependencies between the time steps of a data sequence.

The LSTM layer has two states, which are the *hidden state*, also called the *output state*, and the *cell state*. At time t, the hidden state contains the output of the LSTM layer. The cell state contains the information from the previous time step. At each time step, the LSTM layer adds or removes information stored at the previous time step. The layer checks for updates using *gates*.

The weights that an LSTM layer can learn are named *RecurrentWeights* R and *InputWeigths* W and the bias b. The matrices W, R and b are concatenations of the weights W, R and b formulated for each component. These matrices are as follows in *Equation* 7.1:

$$W = \begin{bmatrix} W_i \\ W_f \\ W_g \\ W_o \end{bmatrix}, R = \begin{bmatrix} R_i \\ R_f \\ R_g \\ R_o \end{bmatrix}, b = \begin{bmatrix} b_i \\ b_f \\ b_g \\ b_o \end{bmatrix} \qquad \textit{Equation 7.1}$$

The components that control the cell state are:

- Input gate *(i)* controls the cell state update. The *Equation 7.2* describes input gates components at time step t:

$$i_t = \sigma_g(W_i x_t + R_i h_{t-1} + b_i) \qquad \textit{Equation 7.2}$$

- Forget gate *(f)* controls level of cell state reset *(Equation 7.3):*

$$f_t = \sigma_g(W_f x_t + R_f h_{t-1} + b_f) \qquad \textit{Equation 7.3}$$

- Cell candidate *(g)* adds information to cell state *(Equation 7.4):*

$$g_t = \sigma_c(W_g x_t + R_g h_{t-1} + b_g) \qquad \textit{Equation 7.4}$$

- Output gate *(o)* controls the level of the cell state added to the hidden state *(Equation 7.5):*

$$o_t = \sigma_g(W_o x_t + R_o h_{t-1} + b_o) \qquad \textit{Equation 7.5}$$

σ_g and σ_c denote respectively the gate activation function and the state activation function.

Figure 7.10 is an illustration of the process of gates forgetting, updating, and outputting cell and hidden state at time step t.

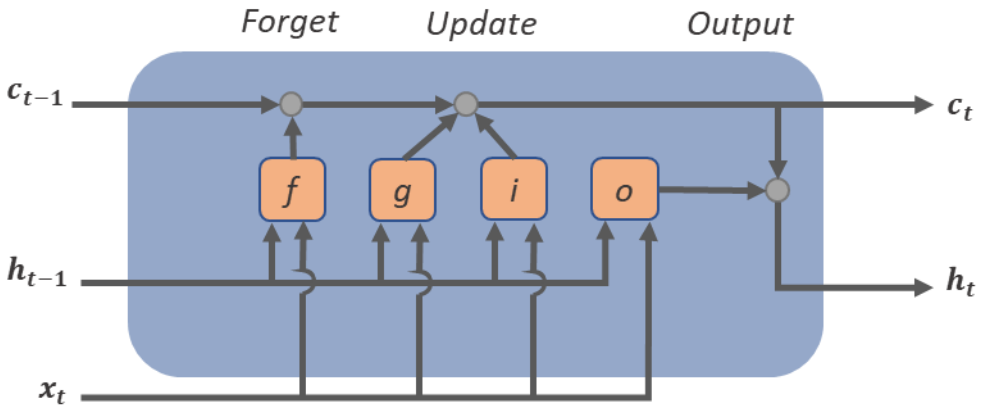


Figure 7.10. Illustration of the process of gates forgetting, updating and outputting cell and hidden state at time step t

7.8.3. Gate Recurrent Unit (GRU)

A GRU layer learn time series dependencies and is well suited, such as LSTM layer, for sequence data.

At each time step t, through *gates*, the hidden state that contain information is updated.

The components that control the cell state are:

- *Reset gate (r)* controls level of state reset. In MATLAB, if ResetGateMode is set to:
    - '*after-multiplication*' or '*before-multiplication*', the *Equation 7.6* describes the components at time steps *t:*

$$r_t = \sigma_g(W_r \boldsymbol{x}_t + b_{Wr} + R_r \boldsymbol{h}_{t-1}) \qquad \textit{Equation 7.6}$$

    - '*recurrent-bias-after-multiplication*', the *Equation 7.7* describes the components at time steps *t*:

$$r_t = \sigma_g(W_r \mathbf{x}_t + b_{Wr} + R_r \mathbf{h}_{t-1}) + b_{Rr} \qquad \textit{Equation 7.7}$$

- *Update gate (z)* controls level of state update. If ResetGateMode is set to:
    - '*after-multiplication*' or '*before-multiplication*', the *Equation 7.8* describes the components at time steps *t*:

$$z_t = \sigma_g(W_z \mathbf{x}_t + b_{Wz} + R_z \mathbf{h}_{t-1}) \qquad \textit{Equation 7.8}$$

    - '*recurrent-bias-after-multiplication*', the *Equation 7.9* describes the components at time steps *t*:

$$z_t = \sigma_g(W_z \mathbf{x}_t + b_{Wz} + R_z \mathbf{h}_{t-1}) + b_{Rz} \qquad \textit{Equation 7.9}$$

- *Candidate state (h)* controls level of update added to hidden state If ResetGateMode is set to:
    - *'after-multiplication'*, the *Equation 7.10* that describes the components at time steps *t*:

$$h_t = \sigma_s(W_{Wh} \mathbf{x}_t + b_{Wh} + r_t \odot (\boldsymbol{R}_h \mathbf{h}_{t-1})) \qquad \textit{Equation 7.10}$$

- '*before-multiplication*', the E*quation 7.11* describes the components at time steps t:

$$h_t = \sigma_s(W_{Wh}\mathbf{x}_t + b_{Wh} + \boldsymbol{R}_h r_t(r_t \odot \mathbf{h}_{t-1})) \qquad \textit{Equation 7.11}$$

- '*recurrent-bias-after-multiplication*', the *Equation 7.12* describes the components at time steps t:

$$h_t = \sigma_s(W_{Wh}\mathbf{x}_t + b_{Wh} + r_t \odot (\boldsymbol{R}_h\mathbf{h}_{t-1} + b_{Rh})) \qquad \textit{Equation 7.12}$$

The weights that a GRU layer can learn are named *RecurrentWeights* R and *InputWeigths* W and bias b. The matrices W and R are concatenations of the weights W and R formulated for each component. These matrices are as follows in *Equation 7.13:*

$$W = \begin{bmatrix} W_r \\ W_z \\ W_h \end{bmatrix}, R = \begin{bmatrix} R_r \\ R_z \\ R_h \end{bmatrix} \qquad \textit{Equation 7.13}$$

Bias b depends on the *ResetGateMode* property.

If ResetGateMode is *'after-multiplication'* or *'before-multiplication'* the bias vector is a concatenation of three vectors *(Equation 7.14)*:

$$b = \begin{bmatrix} b_{Wr} \\ b_{Wz} \\ b_{Wh} \end{bmatrix} \qquad \textit{Equation 7.14}$$

If ResetGateMode is 'recurrent-bias-after-multiplication' the bias vector is a concatenation of six vectors *(Equation 7.15)*:

$$b = \begin{bmatrix} b_{Wr} \\ b_{Wz} \\ b_{Wh} \\ b_{Rr} \\ b_{Rz} \\ b_{Rh} \end{bmatrix} \qquad \textit{Equation 7.15}$$

R indicates that the bias corresponding to the recurrent weights multiplication.

σ_g and σ_s indicates the gate and state activation functions, respectively.

For the hidden state, the *Equation 7.16* describes the components at time steps t:

$$\mathbf{h}_t = (1 - z_t) \odot h_t + z_t \odot \mathbf{h}_{t-1} \qquad \textit{Equation 7.16}$$

7.8.4. Hybrid Deep Learning (hDL) architecture

The inputs of the network are MI-EEG signals with spatial and temporal characteristics. The channels are arranged along rows and the time samples are arranged along the columns. Furthermore, the channels represent the spatial characteristics, and the columns represent the temporal characteristics. The implementation of the network should take this difference into account.

To classify the two types of MI, a hybrid CNN-RNN model was implemented, based on the study [59]. However, the network was modified according to the needs.

In this thesis work, the implemented hybrid network consists of two main blocks. The first is a convolutional block, which contains layers that apply convolution along the rows of the input matrix and along the columns of the feature extracted from the two previous layers.

The convolutional block has the role of extracting the spatial and temporal characteristics of the signal. Between one convolution layer and the next, pool layers are adopted to decrease the amount of data

After the convolution block, the extracted features are fed to a recurrent block, which consist of time recurrent layers that are suitable for extracting and predicting time series data.

Finally, a Classification layer outputs the classification results of the network.

The hybrid network is illustrated, in a general way, in the *Figure 7.11.* The convolution block is shown in yellow, and the orange is used to represent the RNN block.

The convolutional block consists of three convolutional layers that consider the difference between the two dimensions:

- The first is a convolution layer of dimension $N \times 4$. It performs a convolution along the spatial dimension of the input data. The number of convolution kernels is 10
- The second is a convolution layer of dimension $[1 \times 8]$. It performs a temporal convolution of features extracted from the first convolution layer. The number of convolution kernels is 40.
- The thirth is a convolution layer of dimension $[1 \times 4]$. It performs a temporal convolution of features extracted from the previous convolutional layer. The main purpose is to further extract features. The number of convolution kernels is 80.

The second block consists of an LSTM layer with 50 hidden neurons and one GRU layer with 200 neurons. The LSTM layer works well with time sequences. The choice of the GRU layer was made to facilitate a faster network learning process. Their purpose is to classify temporal features extracted from the convolutional block.

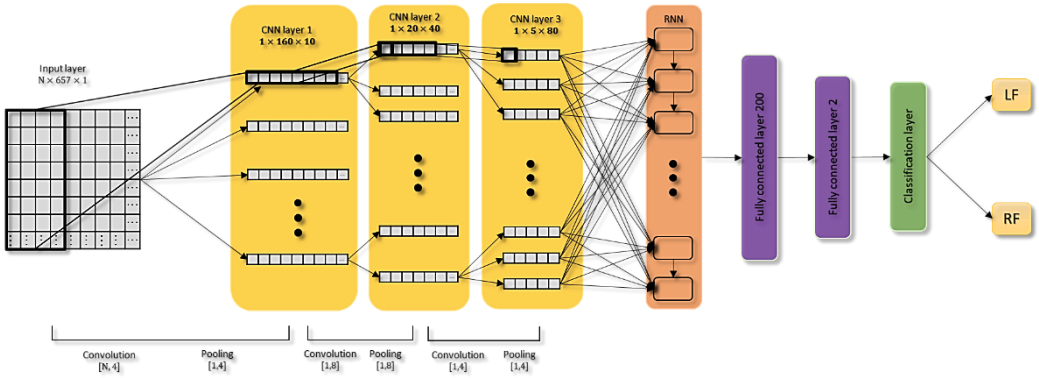


Figure 7.11. Illustration of the hybrid network. The convolution block is shown in yellow, while the RNN block in orange.

Each convolution layer is followed by a Batch Normalization layer (BN layer), an activation layer (ReLU layer) and a Max Pool layer (MP layer).

- Batch normalization layer normalizes mini batches of data across all observations. Specifically, batch normalizes the x_i elements of the input by calculating the mean μ_B and variance σ_B^2, as explained in *Equation 7.17:*

$$x_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \qquad \textit{Equation 7.17}$$

  ϵ is a constant that improves numerical stability when the variance is low [60].

- A ReLU layer performs a threshold operation [61]. Input element less than zero has been set to zero, as follows in *Equation 7.18:*

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \qquad \textit{Equation 7.18}$$

- A max pooling layer performs subsampling by dividing the input into small regions and considering the maximum for each region created [62].

In addition, after the LSTM block, there are two fully connected layers (FC) consisting of 200 neurons and 2 neurons, respectively. Each FC connects all the neurons of the previous layers. In addition, FC multiplies the inputs with a weight matrix and adds a bias vector [63]. The fully connected layer has the purpose of combining all the characteristics extracted from the previous layers.

The last FC layer has two neurons, which correspond to the number of output classes.

After the fully connected layers, there is another activation function, which is called softmax function. In a classification problem, it is generally placed before the classification layer. If the problem requires the classification of two classes, such as in the present case, the softmax function follows *Equation 7.19:*

$$P(c_r|x,\theta) = \frac{P(x,\theta|c_r)P(c_r)}{\sum_{j=1}^{k} P(x,\theta|c_j)P(c_j)} = \frac{exp(a_r(x,\theta))}{\sum_{j=1}^{k} exp(a_j(x,\theta))} \qquad \textit{Equation 7.19}$$

Where:

- $0 \leq P(x,\theta|c_r) \leq 1$ and $\sum_{j=1}^{k} P(x,\theta|c_r) = 1$
- $a_r = ln(P(x,\theta|c_r)P(c_r)), P(x,\theta|c_r)$ is the conditional probability of the sample given class r
- $P(c_r)$ is the class prior probability.

Finally, a classification layer was used. Its purpose is to calculate the cross-entropy loss (loss function) for classification.

The classification layer assigns each input to a class K using the cross-entropy function.

Equation 7.20 represents the loss function, which in this case is a cross-entropy loss function:

$$loss = -\frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{K} w_i t_{ni} \ln y_{ni} \qquad \textit{Equation 7.20}$$

Where:

- N is the number of samples
- K is the number of classes
- w_i is the weight for class i
- t_{ni} is an indicator that the nth sample is a member of the class ith
- y_{ni} is the probability that the network assigns the nth input to a class i [64].

In addition, to solve the overfitting problem, a dropout layer was added at the end of the convolution block and at the end of the LSTM block.

Dropout layer randomly sets all input elements to zero with a probability of 25%. However, as max pooling layer, it has no learning capability.

Figure 7.12 shows a detailed graphical representation of the CNN-RNN network and its connections.

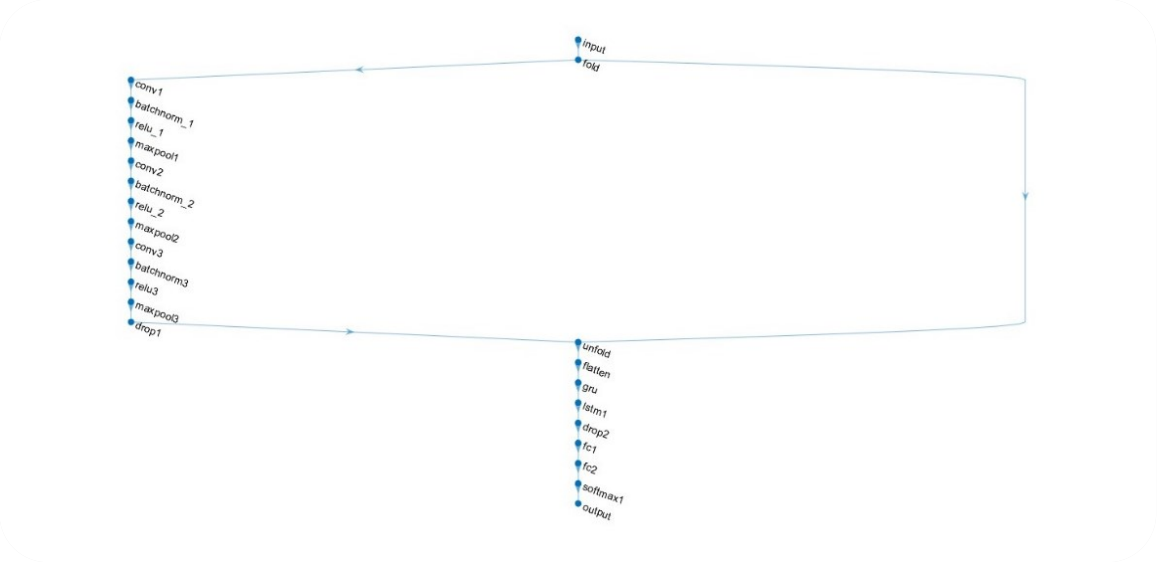


Figure 7.12. Graphical representation of the CNN-LSTM network.

7.9 Evaluation performance

7.9.1. Leave N Subjects Out

When a model is evaluated, it is important to assess its level of generalisation, which means to consider how a model classifies data on which it has not been trained. The data that is used to train the model is called the training set. Whereas, if the model's ability to generalise is to be assessed, the same model must be tested on a test set. The method used to test the performance of the network is the Leave-N-Subjects Out, where N is the number of groups. This method attempts to emulate what happens in real situations, where the network is trained on the data of some subjects and is run on other subjects. Finally, it is a suitable method to assess the level of generalization of the network.

The approach of the method used is very similar to k-fold cross-validation but is more penalizing in terms of results. In fact, unlike k-fold cross-

validation, with the use of Leave N Subjects Out, the validation of the model realized by the network takes place on subjects who were not seen in the training phase. As the EEG is a signal with high variability between different subjects and between trials within the same subject, the performance of the model will be inferior. In fact, the subjects were divided into 5 groups with 80 % of the subjects for the Training and 20 % of the subjects for the Testing set. At each iteration, the network is trained and tested on different groups of subjects. *Figure 7.13* shows how the network is trained and tested. In particular, the subjects that at iteration *i* were part of such Testing set, at iteration *i+1* are considered as Training subjects. The iteration continues until the maximum number of groups considered is reached. When this happens, all subjects in the dataset are taken as both Testing and Training network subjects. Finally, the overall accuracy of the model is the average of the accuracy values obtained for each group.

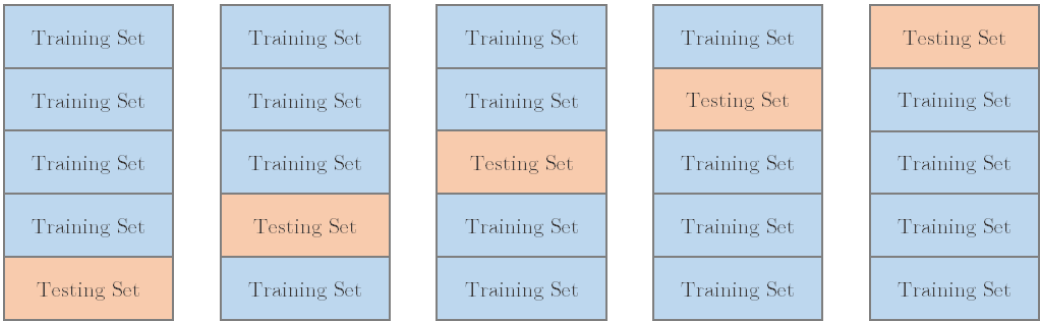


Figure 7.13. Illustration of Leave N Subjects Out groups. Training sets are made with 80% of the total subjects and Testing set are made with 20% of the total subjects.

7.9.2. Precision, Recall and F1 score

Precision, recall and F1 score were evaluated in the following way.

Precision, recall and F1 score, in a classification task, should be applied separately for each of the existing classes.

This means that, for assessing the ability to distinguish left-handed movement from right-handed movement, the LF label was taken as the true positive value and the RF as the false value.

On the other hand, for evaluating the ability to distinguish the imagery movement of the right hand from the left hand, the label RF is considered as a positive value, while the label LF as a false value.

Thus, the values used for the metrics are:

- True Positive (TP), where the classifier recognises the imagery movement of the left fist (or right fist) and the label associated with that motor imagery is LF (or RF),
- False Positive (FP), where the classifier recognises the imagery movement of the left fist (or right fist), but the label is RF (or LF),
- True Negative (TN), where the classifier recognises the imagery movement of the right fist (or left fist) and the label is RF (or LF),
- False Negative (FN), where the classifier recognises the imagery movement of the right fist (or left fist), but the correct label is LF (or RF).

Applying these considerations to *Equations 7.21*, *Equations 7.22,* and *Equations 7.23*, we obtain the results explained in the 'Results' chapter.

$$Precision = \frac{TP}{TP + FP} \quad \text{Equation 7.21}$$

$$Recall = \frac{TP}{TP+FN} \quad \text{Equation 7.22}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad \text{Equation 7.23}$$

8 Results

In this chapter, the results of the proposed model are outlined.

Using a laptop with an Nvidia GeForce GTX 1080 Max-Q graphics card with 4 GB RAM, 8 GB RAM and Intel Core(TM) i7-8565U CPU @ 1.80GHz 1.99 GHz, the code was run a number of times equal to the number of channel groups. The performance of the model, for each group, was validated according to the Leave N Subjects Out method. *Table 8.1* shows the accuracy, precision, recall, F1 score and time required for training.

Groups	Number of channels	Accuracy (%)(mean ± std)	Precision (%)		Recall (%)		F1-score (%)		Time (min)
			LF	RF	LF	RF	LF	RF	
A	3	65,55 ± 1,51	67	64,27	62,57	68,60	64,71	66,36	13,14
B	6	69,51 ± 1,62	70,07	68,95	69,09	69,93	69,58	69,44	13,26
C	10	72,80 ± 1,57	73,77	71,87	71,54	74,09	72,64	72,96	18,65
D	13	73,77 ± 2,71	74,96	72,63	72,08	75,47	73,49	74,02	23,59

E	18	$75{,}96 \pm 1{,}31$	77,21	74,76	74,25	77,68	75,70	76,19	29,15
F	21	$75{,}61 \pm 0{,}49$	76,68	74,56	74,25	77	75,44	75,77	31,82
G	64	$80{,}65 \pm 0{,}97$	81,28	80,02	80,09	81,21	80,68	80,61	116,30

Table 8.1. Table below shows the mean and the standard deviation of accuracy for the five groups of folds. Furthermore, they are displayed other parameters such as Precision, Recall, F1 score calculated for each groups.

From the following graph (*Figure 8.1*), it can be seen how the 64-electrode configuration provided better classification results, reaching 80.65% accuracy at the expense of training time, which reaches that result in 116,3 minutes (approx. 1 h 56 min). Furthermore, it can be seen that the network achieves acceptable results already with an 18 electrode configuration.

Thus, the 18 electrode configuration (group E) is proposed to be a good balance between performance and training time.

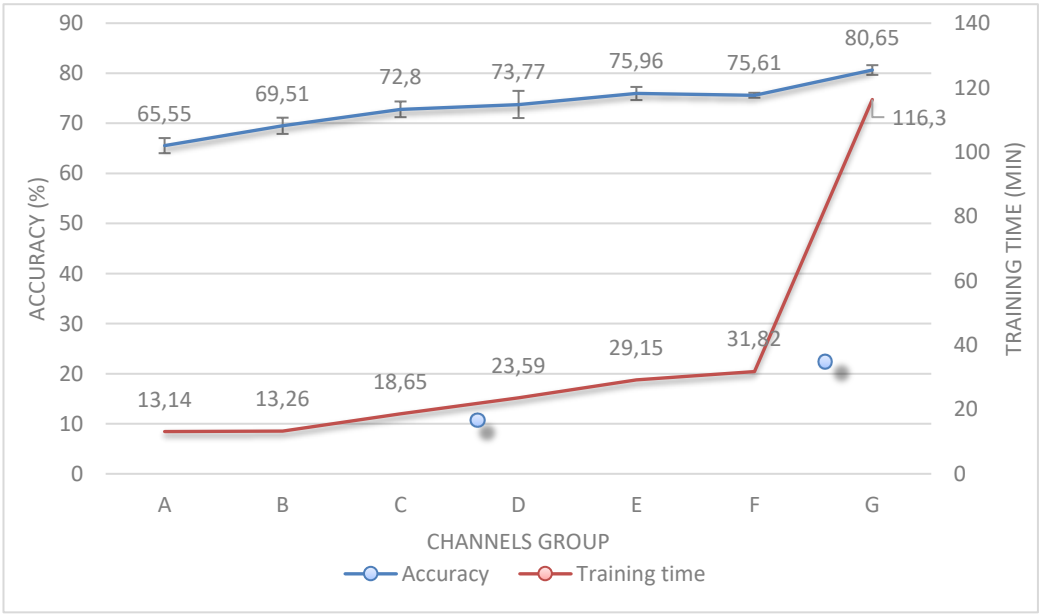


Figure 8.1. The graph displayed shows the accuracy trend as the number of electrodes increase

9 Conclusions

In this thesis work, two motor imagery were classified using an innovative method based on Deep Learning.

For this purpose, a hybrid CNN-RNN network was implemented, consisting of a convolution block and an RNN block. The convolution block consists of three convolutional layers. The first one performs a convolution along the spatial dimension of the signal. The second and third layer perform a convolution along the temporal features previously extracted from the first two layers. The features extracted from the convolution block are fed to the recurrent block formed by the GRU layer and an LSTM layer.

The layers of the recurrent block work very well in the extraction of temporal sequences.

Finally, the output of the LSTM block, passing through fully-connected layers, is fed to a classification layer, which classifies the two imaginary movements, and outputs the classification result.

The network was trained and validated with the Leave N Subjects Out method.

The Leave N Subjects Out method divides the entire number of subjects into five groups. The network is trained on one group of subjects and validated on a group of subjects different from the one used for training. The high variability of the EEG signal between different subjects makes it a challenging method from the point of view of network classification. This method tries to emulate what happens in real applications, where the network is trained on subjects and performed on other subjects.

In addition, a study was performed to evaluate the training time and accuracy as a function of the number of electrodes.

From the study conducted in this thesis, it was possible to observe the ability of Deep Learning in feature extraction. Indeed, the DL network provided good results eliminating time-consuming pre-processing and feature extraction. This aspect represents a good generalisation ability of the network, even though the task is difficult to analyse.

Furthermore, the study on the number of electrodes showed that the 64-electrode configuration leads to excellent results (80.65% accuracy) at the expense of a high training time, but a number of 18 electrodes placed on the SMR cortex, is already a good compromise between training time and accuracy (75.96% accuracy). In addition, a smaller number of electrodes, for a BCI using a motion imagery paradigm, would have the advantage of decreasing the number of electrodes to be placed on the scalp.

As a future implication, it is possible to evaluate the performance of the network in real time. In addition, it is also possible to assess how the results of the network changes by using inputs in a different way, i.e. as scalp maps, each related to the other in terms of time.

10 Bibliography

[1] P. Kolominsky-Rabas, P. Heuschmann, D. Marschall, M. Emmert, N. Baltzer, B. Neundorfer e K. Krobot, «Lifetime cost of ischemic stroke in Germany: Results and national projections from a population-based stroke registry,» *Stroke,*, pp. 1179-1183, 2006.

[2] H. Y. Lee, J. S. Hwang, J. S. Jeng e J. D. Wang, «Quality-adjusted life expectancy (QALE) and loss of qale for patients with ischemic stroke and intracerebral hemorrhage,» *Stroke,* vol. 4, pp. 739-744, 2010.

[3] C. Stinear e W. Byblow, «Predicting and accelerating motor recovery after stroke,» *Current Opinion in Neurology,* vol. 6, pp. 624-630, 2014.

[4] A. J. Jackson, «Neural interfaces for the brain and spinal cord-restoring motor function.,» *Nature Reviews Neurology,* vol. 8, n. 12, pp. 690-699, 2012.

[5] L. Mesin, Introduction to Biomedical Signal Processing, Torino: ILMIOLIBRO, 2017, pp. 223-230.

[6] W.J.Germann e C. Stanfield, Principles of Human Physiology, 6 a cura di, Pearson, Benjamin Cummings, 2005.

[7] C. Giménez, «Composition and structure of the neuronal membrane: molecular basis of its physiology and pathology,» *Revista de Neurologia,* vol. 150, n. 26, pp. 232-239, February 1998.

[8] R. Fields, A. Araque, H. Johansen-Berg, L. S.S., G. Lynch, N. K.A., M. Nedergaard, R. Perez, T. Sejnowski e H. Wake, «Glial Biology in Learning and Cognition,» *Neuroscientist,* vol. 20, n. 5, pp. 426-431, 20 October 2014.

[9] M. Freeman e D. Rowitch, «Evolving concepts of gliogenesis: a look way back and ahead to the next 25 years,» *Neuron,* vol. 80, n. 3, pp. 613-623, 30 October 2013.

[10] M. Y. K. G. Gehrmann J., «Microglia: intrinsic immuneffector cell of the brain,» *Brain Research,* vol. 20, n. 3, pp. 269-287, March 1995.

[11] H. E. Fritsch G, «Electric excitability of the cerebrum (Über die elektrische Erregbarkeit des Grosshirns),» pp. 123-130, 15 June 2009.

[12] P. W. e W. K., «THE SUPPLEMENTARY MOTOR AREA OF THE CEREBRAL CORTEX: A Clinical and Experimental Study,» *AMA Arch NeurPsych.,* vol. 66, n. 3, pp. 289-317, September 1951.

[13] W. C.N., S. P.H., M. D.R., S. W., P. H. T. e T. A.M., «Patterns of localization in precentral and supplementary motor areas and their relation to the concept of a premotor area,» *Res Publ Assoc Res Nerv Ment Dis,* vol. 30, pp. 238-264, 1952.

[14] D. M. Jimsheleishvili S, «Neuroanatomy, Cerebellum,» Jan 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK538167/. [Consultato il giorno 29 August 2022].

[15] N. H. Lee E. Miller, Brain–Computer Interfaces: Principles and Practice, Oxford University Press, 2012, pp. 36-43.

[16] W. J. Daly JJ, «Brain-computer interfaces in neurological rehabilitation,» *Lancet Neurol.,* vol. 7, n. 11, pp. 1032-1043, November 2008.

[17] C.-H. Im, «Basics of EEG: Generation, Acquisition,».*Computational EEG Analysis, Biological and Medical Physics,.*

[18] O. P., «Neurophysiologic basis of EEG,» *J Clin Neurophysiologic,* vol. 3, pp. 186-189, 23 June 2006.

[19] S. JN e M. PJ, «The utility of EEG band power analysis in the study of infancy and early childhood,» vol. 37, n. 3, pp. 253-273, 2012.

[20] «Diagnosis of Alzheimer's disease with Electroencephalography in a differential framework,» *PLos One,* vol. 13, n. 3, 20 Mar 2018.

[21] L. M. N. P. e A. S. Abdullah, «Epilepsy EEG,» *StatPearls [Internet],* May 2022.

[22] W. He, G. C. McConnell e R. V. Bellamkonda, «Nanoscale laminin coating modulates cortical scarring response around implanted silicon microelectrode arrays,» *Journal of Neural Engineering,* vol. 3, n. 4, pp. 316-326, 2006.

[23] S. Breit, J. B. Schulz e A. L. Benabid, «Deep Brain Stimulation,» *Cell and Tissue Research,* vol. 318, n. 1, pp. 275-288, 2004.

[24] K. C. Cheung, «Implantable microscale neural interfaces,» *Biomedical Microdevices,* vol. 9, n. 6, pp. 923-938, 2007.

[25] A. B.Ajiboye e R. F.Kirsch, «Chapter 27 - Invasive Brain–Computer Interfaces for Functional Restoration,» in *Neuromodulation*, Elliot S. Krames, P. Hunter Peckham, Ali R. Rezai, 2018, pp. 379-391.

[26] J. L. Roland, W. Z. Ray e E. C. Leuthardt, «Neuroprosthetics,» in *Youmans and Winn Neurological Surgery, Seventh Edition*, pp. 754-758.

[27] M. Teplan, «FUNDAMENTALS OF EEG MEASUREMENT,» *MEASUREMENT SCIENCE REVIEW,* vol. 2, n. 2, 2002.

[28] C. MD e P. J., «P3a and P3b from typical auditory and visual stimuli,» *Clin Neurophysiol.,* vol. 110, n. 1, pp. 24-30, January 1999.

[29] J. Polich e C. Margala, «P300 and probability: comparison of oddball and single-stimulus paradigms,,» *International Journal of Psychophysiology,,* vol. 25, n. 2, pp. 169-176, 1997.

[30] G. McCarthy e E. Donchin, «A metric for thought: a comparison of P300 latency and reaction time,» *Science,* vol. 211, n. 4477, pp. 77-80, 2 January 1981.

[31] M. Kutas, G. McCarthy e E. Donchin, «Augmenting Mental Chronometry: The P300 as a Measure of Stimulus Evaluation Time,» *Science,* vol. 4305, pp. 792-795, August 1977.

[32] F. Piccione, F. Giorgi, P. Tonin, K. Priftis, S. Giove, S. Silvoni, G. Palmas e F. Beverina, «P300-based brain computer interface: reliability and performance in healthy and paralysed participants,» *Clin Neurophysiol.,* vol. 117, n. 3, pp. 531-537, Mar 2006.

[33] C. Bell, P. Shenoy, R. Chalodhorn e R. Rao, «Control of a humanoid robot by a noninvasive brain-computer interface in humans,» *J Neural Eng.,* vol. 5, n. 2, pp. 214-220, June 2008.

[34] I. Iturrate, J. M. Antelis, A. Kubler e J. Minguez, «A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation,» *IEEE Trans. Rob.,* p. 614–627, 2009.

[35] G. Pfurtscheller, C. Brunner, A. Schlögl e L. d. S. F.H., «Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks,» *Neuroimage,* vol. 31, n. 1, pp. 153-159, 15 May 2006.

[36] C.-N. J., C.-E. R.I., C.-M. P., E.-V. D. e G.-M. J., «Motor Imagery-Based Brain-Computer Interface Coupled to a Robotic Hand Orthosis Aimed for Neurorehabilitation of Stroke Patients,» *J Healthc Eng,* 3 April 2018.

[37] E. López-Larraz, A. Sarasola-Sanz, N. Irastorza-Landa, N. Birbaumer e A. Ramos-Murguialday, «Brain-machine interfaces for rehabilitation in stroke: A review. NeuroRehabilitation,» vol. 43, n. 1, pp. 77-97, 2018.

[38] G. Celesia e N. S. Peachey, «Visual evoked potentials and electroretinograms,» *Electroencephalography-Basic principles, clinical applications and related fields,* pp. 1017-1043, 2005.

[39] B. Allison, C. Brunner, V. Kaiser, G. Müller-Putz, C. Neuper e G. Pfurtscheller, «Toward a hybrid brain-computer interface based on imagined movement and visual attention,» *Journal of Neural Engineering,* 2010.

[40] L. J. Trejo, R. Rospial e B. Matthews, «Brain-computer interfaces for 1-D and 2-D cursor control: designs using volitional control of the EEG spectrum or steady-state visual evoked potentials,» *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* pp. 225-229, 2006.

[41] J. Faller, G. R. Müller-Putz, D. Schmalstieg e G. Pfurtscheller, «An Application Framework for Controlling an Avatar in a Desktop-Based Virtual Environment via a Software SSVEP Brain–Computer Interface,» *Presence Teleoperators & Virtual Environments,* vol. 19, n. 1, pp. 25-34, February 2010.

[42] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey e A. e. a. .. Kübler, «A spelling device for the paralysed,» *Nature,* p. 297 – 298, 1999.

[43] J. Colebatch, «Bereitschaft spotential and movement-related potentials:,» *M ovement Disorders ,,* pp. 601-610, 2007.

[44] H. Shibasaki e M. Hallett, «What is the Bereitschaft spotential?,» *Clinical Neurophysiology,* pp. 2341-2356, 2006.

[45] R. Beisteiner, P. Höllinger, G. Lindinger, W. Lang e A. Berthoz, «Mental representations of movements . Brain potentials associated with imagination of hand movements,» *Electroencephalography and Clinical,* p. 83 – 193, 1995.

[46] A. Kübler, N. Neumann, J. Kaiser, B. Kotchoubey, T. Hinterberger e N. .. Birbaumer, «Brain-computer communication: self-regulation of slow cortical potentials for verbal communication,» *Archives of Physical Medicine and Rehabilitation,* p. 1533 – 1539.

[47] C. Neuper e G. Pfurtscheller, «Brain-computer interfaces: non-invasive and invasive technologies,» *Brain-Computer Interfaces: Non-invasive and Invasive Technologies,* pp. 65-78, 2010.

[48] A. A. L. G. L. H. J. I. P. C. M. R. .. &. S. H. E. Goldberger, «PhysioBank, PhysioToolkit, and PhysioNet: Components of a new

research resource for complex physiologic signals,» *Circulation [Online],* vol. 23, p. e215–e220, 2020.

[49] I. The Mathworks, «MATLAB version 9.12.0.2039608 (R2022a) Update 5». 2022.

[50] D. Hagemann e E. Naumann, «The effects of ocular artifacts on (lateralized) broadband power in the EEG,» *Clin Neurophysiol.,* vol. 112, n. 2, pp. 215-231, February 2001.

[51] K. I., «A Brief Summary of EEG Artifact Handling in Brain-Computer Interface,» *IntechOpen, 2021 [Online],* 2021.

[52] M. Y., «Removal methods of EMG Artifacts from EEG,» *J. Phys.: Conf. Ser,* 2021.

[53] Y. LeCun, Y. Bengio e G. Hinton, «Deep Learning,» *Nature,* vol. 521, pp. 436-444, 2015.

[54] M. Clerc, L. Bougrain e F. Lotte, «Brain Computer Interfaces,» in *Brain Computer Interfaces: Foundations and Methods*, New York, Clerc, M,; Bougrain, L.; Lotte, F., 2016.

[55] W. Fang, P. Love, H. Luo e L. Ding, «Computer vision for behaviour-based safety in construction: a review and future,» *Adv Eng Inform,* 2020.

[56] D. Palaz, M. Magimai Doss e R. Collobert, «End to end acoustic modeling using convolutional neural networks for HMM-based automatic speech recognition,» *Speech Community,* vol. 108, pp. 15-32, 2019.

[57] H. Li, Z. Deng e H. Chiang, «Lightweight and resource constrained learning network for face recognition with performance optimization,» *Sensors,* vol. 20, n. 21, 2020.

[58] MATLAB, «convolution2dLayer,» MathWorks, [Online]. Available: https://it.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.convolution2dlayer.html;jsessionid=493e781ecaa0e11615e572288eaf. [Consultato il giorno 30 September 2022].

[59] K. Zhu, S. Wang, D. Zheng e D. Mengxi, «Study on the effect of different electrodechannel combinations of motor imagery EEGsignals on classification accuracy,» *IET journals,* 2018.

[60] I. The MathWorks, «batchNormalizationLayer,» The MathWorks, Inc., [Online]. Available: https://it.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.batchnormalizationlayer.html#d124e22266. [Consultato il giorno 11 October 2022].

[61] I. The MathWorks, «reluLayer,» The MathWorks, Inc., [Online]. Available: https://it.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.relulayer.html. [Consultato il giorno 11 October 2022].

[62] I. The MathWorks, «maxPooling2dLayer,» The MathWorks, Inc., [Online]. Available: https://it.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.maxpooling2dlayer.html. [Consultato il giorno 11 October 2022].

[63] I. The Mathworks, «FullyConnectedLayer,» The Mathworks, Inc., [Online]. Available: https://it.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.fullyconnectedlayer.html#d124e72791. [Consultato il giorno 11 October 2022].

[64] I. The MathWorks, «classificationLayer,» The MathWorks, Inc., [Online]. Available: https://it.mathworks.com/help/deeplearning/ref/classificationlayer.html#bu80saf. [Consultato il giorno 11 October 2022].

[65] H. Gollee, I. Volosyak, A. McLachlan, K. Hunt e A. Gräser, «An SSVEP-based brain-computer interface for the control of functional electrical stimulation,» *IEEE Trans Biomed Eng.,* vol. 58, n. 8, pp. 1847-1855, August 2010.

[66] M. Shanker, M. Hu e M. Hung, «Effect of data standardization on neural network training,» vol. 24, n. 4, pp. 385-397, 1996.