

UNIVERSITÀ DEGLI STUDI DELL'AQUILA

Department of Information Engineering, Computer Science, and Mathematics

Doctoral Program in Information and Communication Technology
Emerging computing models, software architectures, and intelligent systems

GSD: 01/INFO-01

SSD: INFO-01/A

Advanced Computational Models for Neuroinformatics and Brain-Computer Interfaces

Thesis Advisor
Prof. Filippo Mignosi

Thesis Co-Advisor
Prof. Giuseppe Placidi

Daniele Lozzi
281734

Academic Year 2023 - 2024

Contents

Abstract **iv**

1 Introduction **1**

I Materials and Methods **4**

2 The Brain **5**
2.1 Brain Physiology 5
2.2 Brain Functions 13

3 Brain Activity Measurement and Analysis **23**
3.1 Neuroimaging techniques as a whole 23
3.2 Electroencephalography (EEG) 27
3.3 Computational tools for EEG 32
3.4 EEG-based Brain-Computer Interfaces 49

II Specific Studies and Related Computational Models **55**

4 Motor Execution **56**
4.1 Analysis of the EEG Resting-State Signals for BCI 59
4.2 Deep Learning Architecture analysis for EEG-Based BCI 63
4.3 The MOVING dataset 70

5 Sleep **87**
5.1 Materials and Methods 90
5.2 Descriptive Statistics 98
5.3 Results 98

6 Emotions **99**
6.1 The challenge of Emotion Recognition 102
6.2 Brain Functional Connectivity of Emotions 116
6.3 4D LSTM Neural Network for “relax” emotional state classification 121
6.4 Neural correlations of visual emotional processing 129
6.5 REMEMO Dataset 147
6.6 Discussion 155

III Conclusions **157**

7 Conclusive Hints and future developments **158**

IV Bibliography **160**

Bibliography **161**

V Appendix **187**

A Sleep Stage Distribution for Chapter 5 **188**

B Psychological Questionnaires used in Sections 6.4 and 6.5 **198**

C Scoring Psychological Questionnaires for Chapter 6.4 **200**

D Code for Section 6.3 **205**
D.1 Interpolation map 205
D.2 Data augmentation 207
D.3 Extraction min set of channels 208

E Equations **211**

VI Lists **213**

List of Figures **214**

List of Tables **220**

List of Algorithms **223**

Acronyms **225**

Nomenclature **232**

Candidate's Publications **251**

VII Acknowledgement **254**

Acknowledgement **255**

Abstract

Neuroinformatics is an emerging interdisciplinary field aiming to develop methods, algorithms, and visualisation tools for the study of the brain and nervous system. This discipline intersects the field of brain-computer interfaces (BCIs), whose purpose is to exploit external signals causing detectable physiological reactions in the brain or to recognise signals produced by aware activities in the brain as an alternative communication output toward the external world. This thesis uses electroencephalography (EEG) to measure the activity of the human cerebral cortex related to specific brain functions: movements, sleep, and emotions. In addition, it presents advanced computational models to process, analyse, and decode the resulting signals, which are nonlinear and nonstationary and are greatly affected by noise and artifacts. The final goal is to decode the unknown signals of emotions, one of the most complex neurocognitive domains due to its extreme subjectivity and the impossibility of verifying the labels in the collected data. This is done first by validating the proposed models on data allowing well-known brain functions, such as movements and sleep, and then by applying them to the complex and subjective signals from emotions. An evaluation of the activation protocols used to elicit brain activity related to the above functions is furnished, and specific activation protocols are presented for collecting bias-reduced, specific datasets. Experimental results and comparisons with state-of-the-art models are presented and discussed, highlighting the pros and cons of the proposed strategies and suggesting some hints for future developments.

Keywords: *Neuroinformatics; Brain-Computer Interface; Electroencephalography; Artificial Intelligence; Deep Learning; Signal Processing.*

Chapter 1

Introduction

The thesis presents research carried out in neuroinformatics [1, 2], an emerging interdisciplinary area that deals with the application of mathematical, engineering, and computer sciences to neuroscience and cognitive science in all its subfields. This work proposes effective computational models for dealing with Electroencephalography (EEG) signals related to some brain functions, such as movements, sleep, and emotions, and brain-computer interfaces exploiting the signals from the above functions. The EEG is a measurement tool that allows recording the electrical signals of the brain that come from the cortex.

The final proposal is to enable the extraction of patterns within multimodal datasets and make it possible to analyse and extract the nonlinear relationships and emergent behaviours in complex patterns, such as those related to emotions, starting from those allowing to well-known functions. This is done by presenting advanced computational models suitable for the resulting complex, nonlinear, nonstationary, and noisy EEG signals. The strategy used to validate the proposed models starts from signals whose patterns and neurophysiological mechanisms are known, such as movement and sleep, to decode emotions, a field of neuroscience where the scientific community has not yet reached an agreement on the functioning theories and interpretation of results. The proposed strategies make possible the extraction of patterns within datasets collected using specifically designed protocols to analyse emergent behaviour in complex systems, such as those related to the functions considered in the brain. As shown in Fig. 1.1, computational tools that allow the study of complex signals EEG have been studied in the context of the neurophysiological mechanisms underlying the investigated domains. The models and algorithms proposed therein are mostly Artificial Intelligence (AI)-based since deterministic algorithms have been proven to be unable to analyse the nonlinear and nonstationary nature of EEG, mainly related to the above functions. Despite its nonlinear and nonstationary nature, EEG was chosen as a neuroimaging technique because, unlike other techniques, it can be used in an ecological context, being used not only in a controlled laboratory environment such as functional Magnetic Resonance Imaging (fMRI), but it is also low-cost and easy to wear.

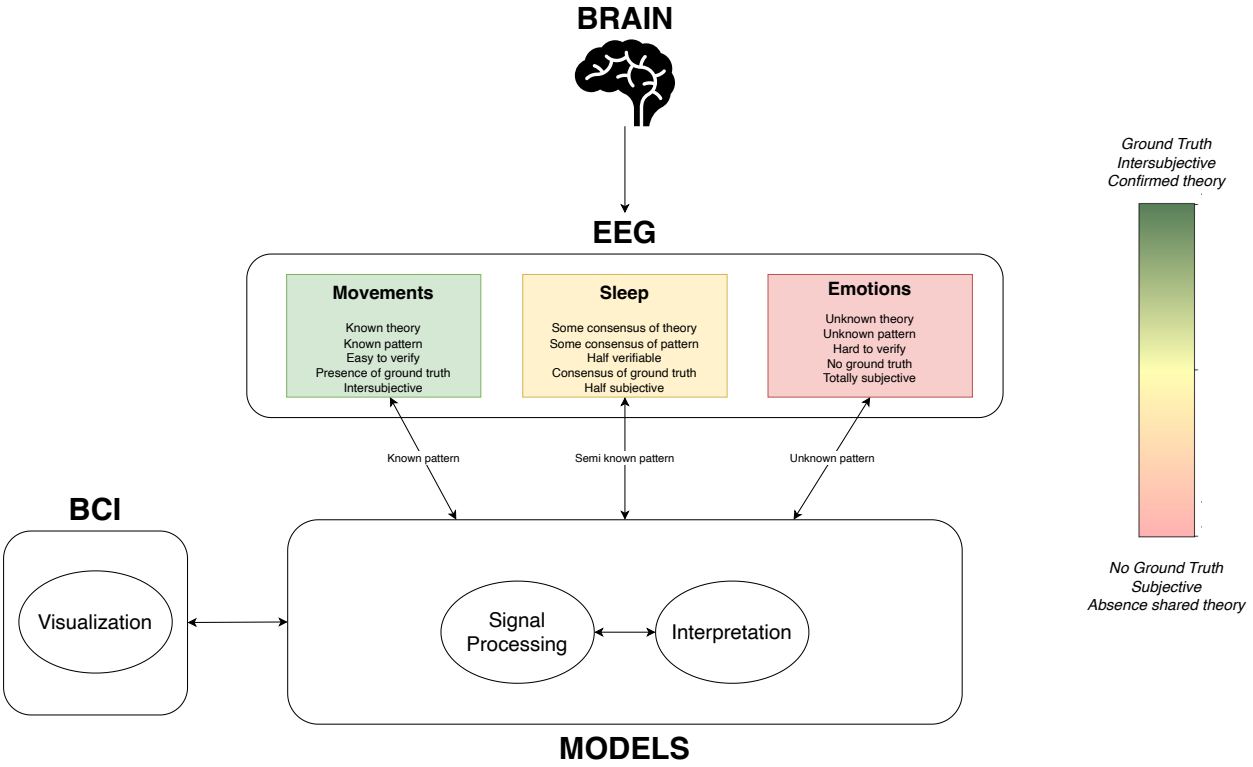


Figure 1.1: This manuscript addresses topics of increasing complexity, both in terms of scientific consensus and the absence of a ground truth. This increasing complexity culminates in the domain of emotions, where it is impossible to have a definite ground truth in emotional datasets. This is due to the high inter-subject variability and the instability of emotions within the same subject, as well as the lack of a universally accepted theory of emotions within the scientific community.

The thesis has a gradual structure. First, it explores the potential of AI to analyse and classify eye movements or sleep stages on brain-computer interfaces. Second, it explores the possibility of transferring the knowledge gained in the first domain to the study of emotions. Specifically, the goal of decoding emotions not only for Brain-Computer Interfaces (BCIs) but also to increase neuroscientific knowledge (one of the goals of neuroinformatics) is pursued by first applying the methods studied to a domain whose neurophysiology is well known, namely Motor Execution (ME), and which can be experimentally verified through direct observation. Subsequently, the techniques to the domain of sleep, whose brain patterns have been well-known for years with numerous experimental confirmations and where only some macroscopic effects can be verified were applied. Finally, based on these studies, investigate the domain of emotions, whose physiology is still poorly understood compared to the previous two, and there is still no broad scientific consensus on its functioning, its circuits, and neural patterns. Moreover, emotions have a very high component of subjectivity and intra-subject instability, unlike sleep and ME.

The Research Questions (RQs) that this work aims to investigate are: (RQ1) What computational models allow for an effective analysis of EEG data in domains whose patterns and neurobiological theory are known, (RQ2) Whether it is possible to translate the models into a domain whose patterns are unknown, such as emotions, and, as a consequence, (RQ3) Whether specific protocols could be defined for the elicitation of emotions with reduced biases.

From an experimental point of view, this thesis presents a series of experimental studies aimed at improving computational systems for analysing EEG data in the field of Affective Computing

(AC). The concept of AC was first introduced by R. Picard [3], who defined it as "computing that relates to arises from or influences emotions".

The thesis is structured as follows. In Part I, the neurophysiological concepts, neuroimaging tools, and neuroinformatics techniques necessary to understand the manuscript are presented. In Part II, all the experimental studies conducted for the scope of the work have been presented and discussed, ordering them according to the increasing complexity of the data: movement, sleep, and emotions. Finally, in Part III, the results of the experimental studies, the limitations and future developments are summarised.

Part I

Materials and Methods

Chapter 2

The Brain

In this chapter, a brief introduction to the anatomy, physiology, and cognitive functions will be provided. The information in this chapter is taken from [4–10]

The brain is part of Central Nervous System (CNS), a complex system composed of numerous levels that governs the whole organism.The brain is the result of gene expression that begins with embryogenesis and lasts a lifetime. The most common analogy between the brain and a computer is that the mind is like the software and the organ is the hardware, which, although inaccurate, is a very common comparison to describe two components of a large system. Among the earliest philosophers who addressed this issue is Descartes, who defined that mind and brain are fundamentally different substances: the brain being physical and the mind being nonphysical or immaterial. This approach is referred to as dualism, as opposed to the monistic approach proposed in particular by neuropsychology, which suggests that the brain is not merely a sensorimotor machine but a complex organ that generates experiences and consciousness. Many modern neuroscientific studies highlight the interplay between the mind and the brain, asserting that mental processes can affect brain function and conversely. Recent findings in psychological therapy and psychotherapy appear to support this perspective [11].

2.1 Brain Physiology

2.1.1 Brain Cells

Brain cells are the basic units that make up CNS, and communication between them is a fundamental part of their functioning. The foundations of modern neuroscience were built by Camillo Golgi [12] and Ramon y Cajal [13], who demonstrated that neurons are discrete entities that communicate through specialized contacts. Subsequently, using imaging techniques, it was possible to study the two main components of the CNS: neurons and glia. In Fig. 2.1, a paint of brain cells made by Cajal is reported.



Figure 2.1: A Cajal's paint representing neurons. Image from the "*Cajal Institute*" and the "*Spanish National Research* Council".

Glia Glial cells support the vital functions of neurons and the system in general, maintain the homeostasis of the CNS and isolate axons through the myelin sheath. They also modulate the speed of propagation of nerve signals. Recently, many studies have confirmed that glia are involved in numerous neural processes, defining the neuron-astrocyte network as the neuron-astrocyte network. In Fig. 2.2 an illustration of glia cells is shown. The glia of CNS are different from those of Peripheral Nervous System (PNS): in the former there are astrocytes that regulate the chemical environment, and while oligodendrocytes cover the axons with myelin, in the latter there are Schwan cells.

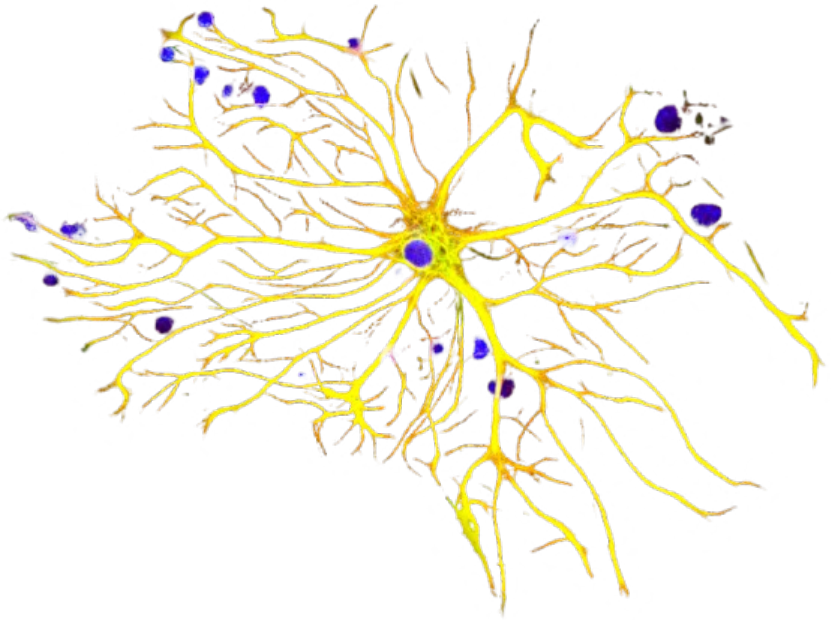


Figure 2.2: An astrocyte. (Author: "GerryShaw", CC-SA).

Neurons Neurons are cells specialized in the transmission of information; they consist of the cell body, axon, dendrites, and create synapses with other neurons. The synapses can be chemical or electrical, depending on the brain area in which the neurons are located, and the signal that triggers communication is given by the exchange of ions with the extracellular environment, which, via the action potential, reaches the dendrites that release the electrochemical message that will propagate to the other neurons in the network, a message that can be excitatory or inhibitory at different intensities. There are many types of neuron based on their functions and shapes: unipolar, bipolar, and multipolar neurons. In Fig. 2.3 a multipolar neuron is shown.
Information processing in CNS occurs with two potentials:

Membrane potential is maintained due to the Na+/K+ pump and the electrical gradient that is generated due to the selectively permeable membrane, which maintains the membrane potential at -70 mV;

Action potential occurs when the sum of excitatory and inhibitory postsynaptic potentials exceeds an activation threshold, opening ion channels that reverse the potential difference from -70 to +35 mV, and which are kept constant thanks to Ranvier's nodes, i.e. breaks in the myelin sheath along the axon. An illustrative diagram is shown in Fig. 2.9.

When this potential reaches the bottom of the dendrites, they release neurotransmitters through terminal buttons that trigger a potential in the neurons connected to it. This conduction and release of neurotransmitters also generate outward currents, which when combined lead to the generation of a signal that is also detectable at the EEG.

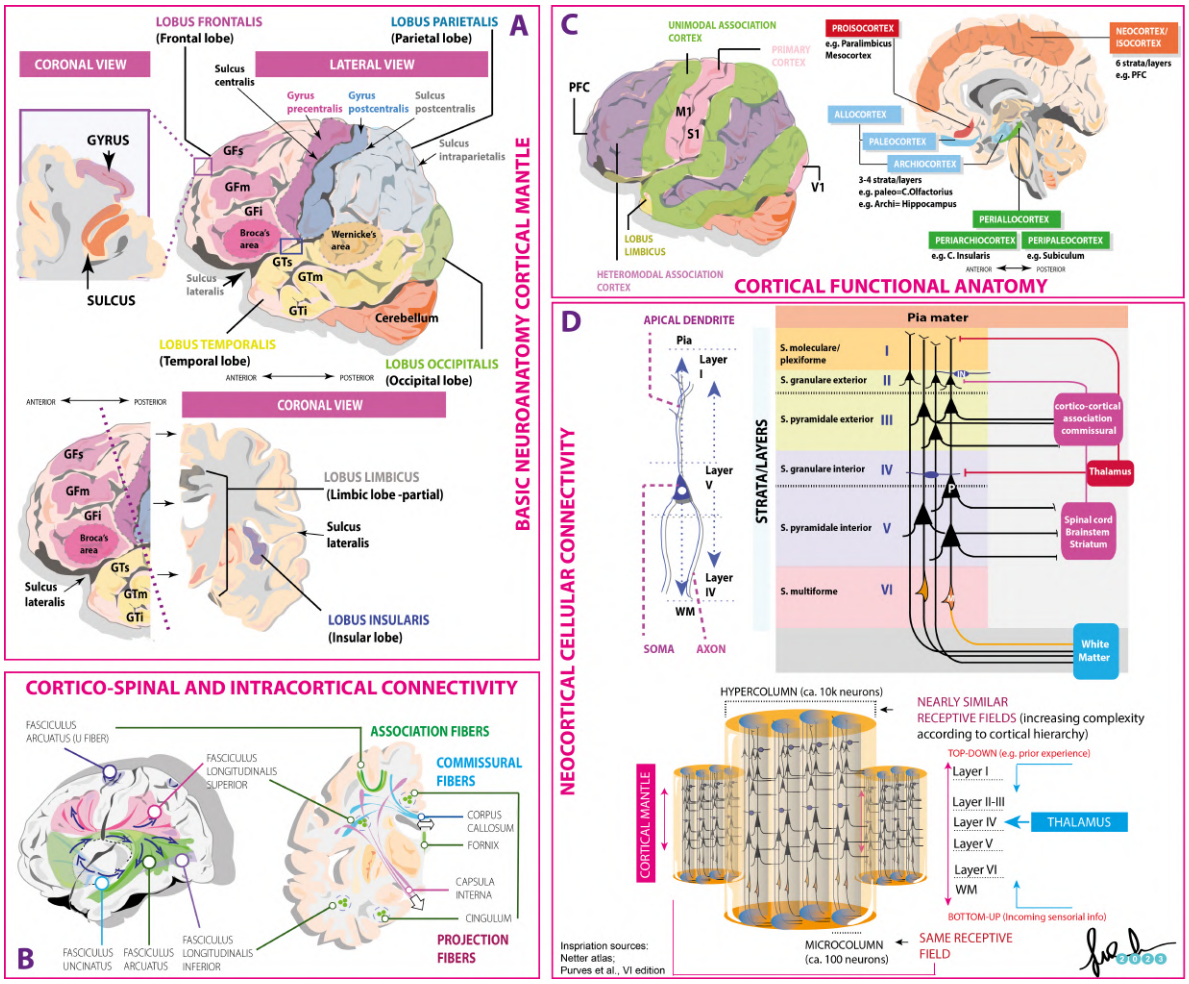


Figure 2.4: An overview of the CNS. Image from [14].

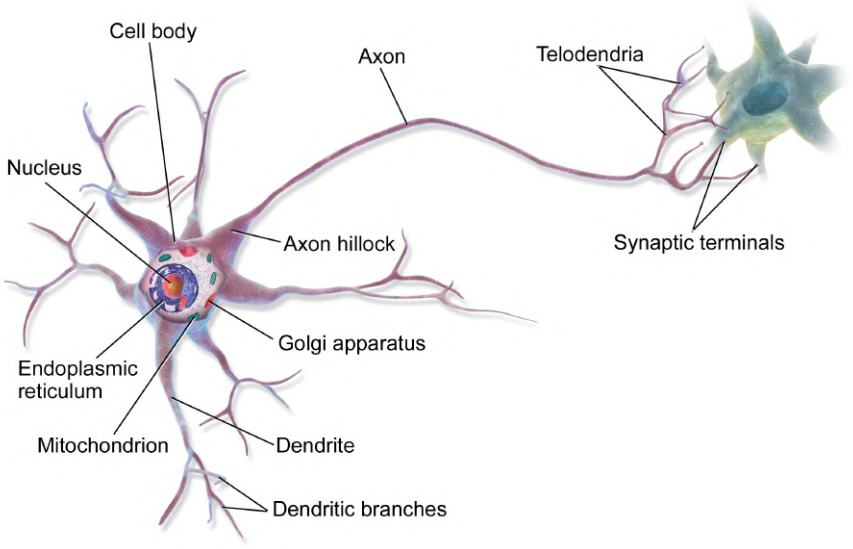


Figure 2.3: Multipolar neuron with its several components. (Author: "BruceBlaus", CC-BY).

2.1.2 Brain structures

The structure of CNS is described in the following with reference to the three anatomical planes: coronal, sagittal, and transversal, as shown in Fig. 2.5.

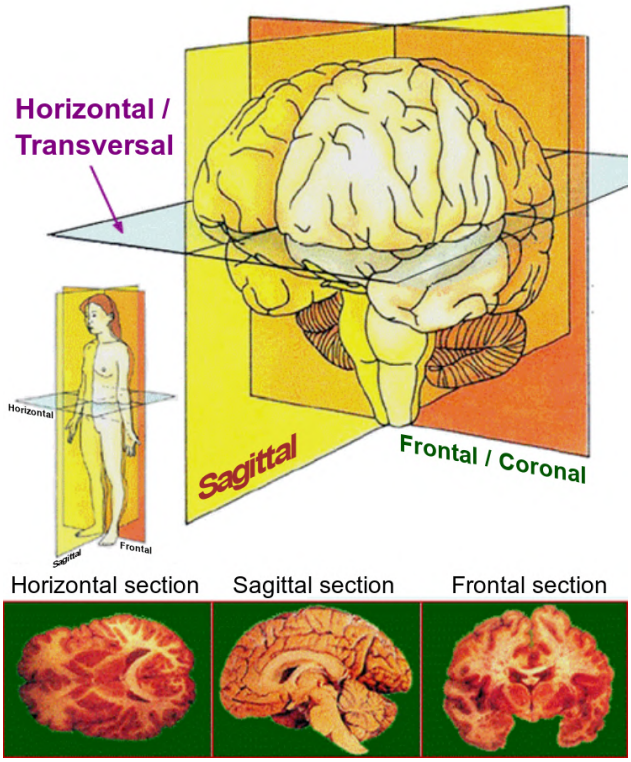


Figure 2.5: The anatomical planes. (Author: "Zwarck", CC-SA).

The CNS can be divided into seven main structures, shown in Fig. 2.4, described below:

Spinal Cord represents the caudal part of the spinal cord and receives information from the muscles, skin, and, in turn, controls the limbs. It contains many clusters of neurons that form the nuclei of the cranial nerves, which also receive information from the head and provide control of the eyes and facial muscles, while others are responsible for analyzing sensory information.

Medulla rostral to the spinal cord is responsible for many vital functions such as breathing.

Pons sends movement information generated by the hemispheres to the cerebellum.

Cerebellum is connected to the brainstem via the pedicles and modulates the force and amplitude of movements and is responsible for programming the movement itself.

Mesencephalon controls the functions of eye movements and senses. The mesencephalon, together with the pons and medulla, is the brain stem.

Diencephalon composed of the thalamus, which analyzes information before it reaches the hemispheres, and the hypothalamus, which regulates information from the endocrine and visceral systems.

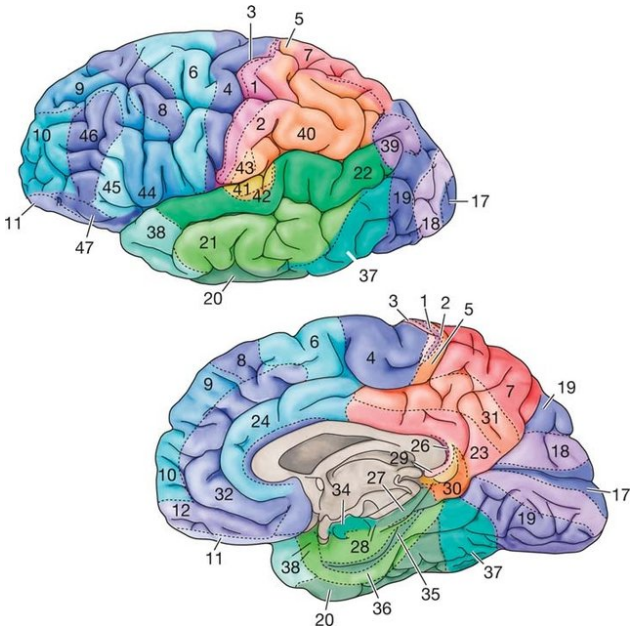


Figure 2.6: The Brodmann areas. Image from [15].

Cerebral Hemispheres formed by the cortex that deals with higher cognitive functions, the basal nuclei that regulate motor performance, the hippocampus that deals with mnestic traces and the amygdala that deals with emotion regulation. Finally, the cortex consists of four lobes for each hemisphere: frontal, parietal, temporal, and occipital. The hemispheres are connected by the corpus callosum, a bundle of axons. The cerebral cortex has regions that have distinct functions, such as movement, analysis of visual input, sound, language, touch, etc., and one of the most widely used categorizations is Brodmann's 1909, shown in Fig. 2.6, and it is organized into seven layers, which are cytologically distinct.

The CNS is a complex system with emergent behavior that can therefore only be appreciated by studying all the levels and properties that it displays. There are five principles that govern the main functions: each functional system is composed of different structures that perform the analysis of information at different levels, which is then transmitted from one structure to another. Within the levels, two macro classes of neurons can be identified: projection neurons, which carry information to the next level, and interneurons, which only establish contact with the local cells involved in processing. Each component of the system under investigation is connected by pathways that can be distinguished anatomically and whose bundles of fibers are identifiable. Furthermore, the organization of the systems is ordered, and precise topographical maps of the system can be identified. Furthermore, the organization of each functional system has a hierarchical structure that, therefore, follows a precise direction. Finally, functional systems are counterlateral, and nerve bundle crossings occur at the spinal level and are called "decussation".

The organization of the nervous system can be divided into sensory systems, motor systems, and association systems that mediate complex brain functions, and is divided into CNS and PNS. The CNS includes the brain and spinal cord, while the PNS includes sensory neurons that connect to receptors. In the CNS, cells are organized in two different ways: there are nuclei, which are clusters and are located in the brainstem, spinal cord, and telencephalon, while the cortices have a layered distribution. Axons, on the other hand, constitute the bundles that connect structures, and those

that cross the sagittal plane are called commissures. Nuclei are generally referred to as gray matter, while bundles are referred to as white matter and are covered by the myelin sheath, whose absence leads to cognitive and physical deficits due to an inability to conduct action potentials. This tissue is attacked during Multiple Sclerosis (MS), and its deterioration determines the progression of the pathology. A characteristic of neural systems is the representation of information in topographic maps, which reflects a point-to-point correspondence of the sensory periphery, while in other neural systems where information is not necessarily represented spatially, computational maps integrate stimulus information and create the sensation that reaches consciousness. The structures of the brain can be divided as follows, as shown in Fig. 2.7.

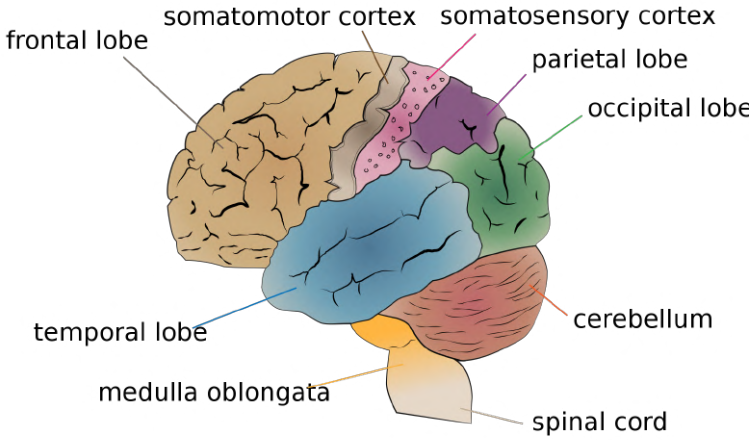


Figure 2.7: The brain structures. Creative Commons license.

A fundamental structure of the cortex is the central sulcus, which separates the frontal lobe from the parietal lobe and houses the topographic representation of the motor system (precentral) and the sensory system (postcentral).

The electrical activity of the brain can be distinguished in two levels: intracellular, given by the exchange of ions between the neuron and the extracellular environment due to the selectively permeable membrane, and the activity of populations of neurons that discharge synchronously, generating a current detectable by instrumentation such as EEG. In particular, the EEG detects the current generated by cortical layers III "external pyramid layer", V "internal pyramid layer" and VI, "polymorphic layer", whose populations discharging together make the electrical signal so high as to be detectable by electrodes. The cortical layers are visible in Fig. 2.8. One problem with EEG is that it is unable to detect currents that are generated in an opposite direction and thus cancel out, as shown by the example in Fig. 2.10.

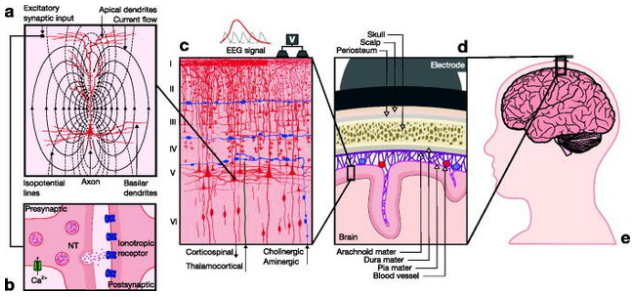


Figure 2.8: Cortex's layers and EEG current generation. Image from [16].

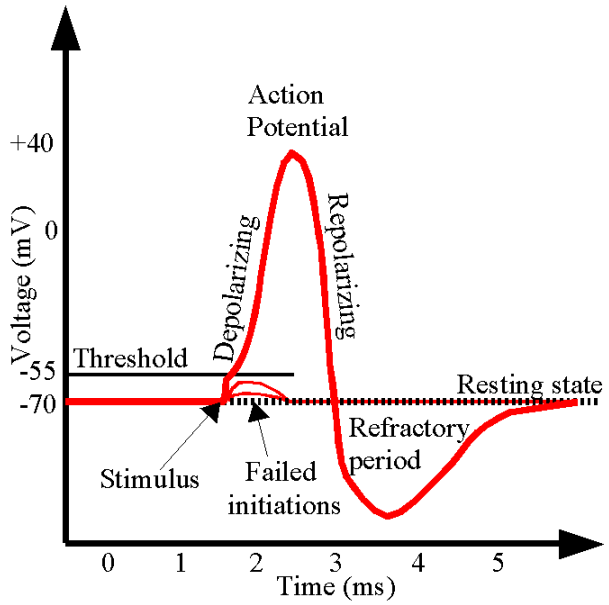


Figure 2.9: Action potential diagram. Image from [17].

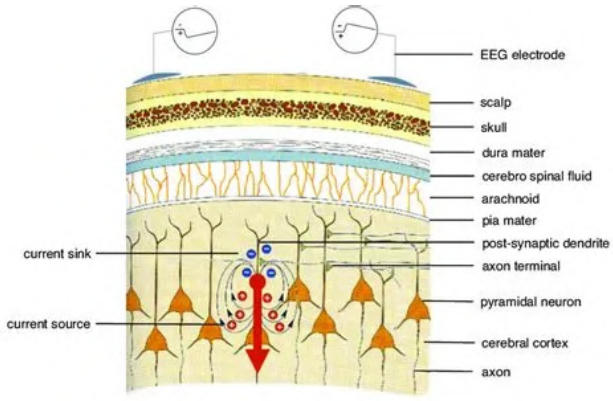


Figure 2.10: Currents detected by EEG. Image from [18].

The EEG is the signal that results from the discharge of neurons from the thalamus, the structure responsible for the integration and transmission of information from the lower centers to the cortex, and it is known that the discharge of the thalamus through the thalamocortical network is fundamental to maintaining the state of awakening. The EEG is the result of the sum of the synap-

tic synchronous potentials of the dendrites of the cortical neurons, which in turn are stimulated by the thalamocortical circuit, visible in Fig. 2.11.

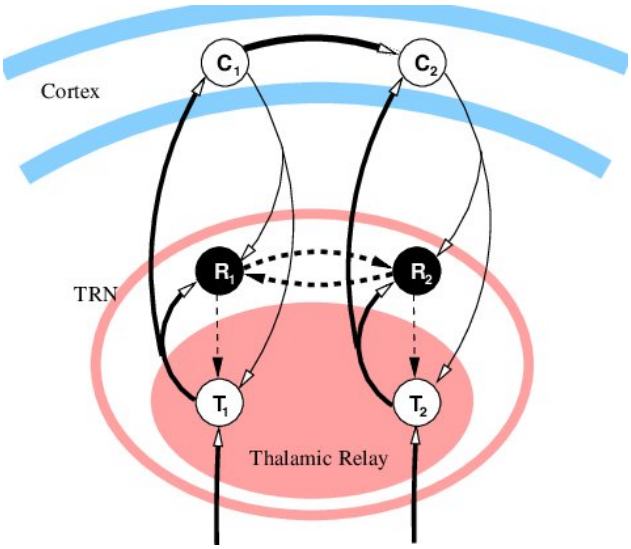


Figure 2.11: Thalamo-cortico-thalamic circuits. Image from [19].

2.2 Brain Functions

The CNS, and in particular the brain, governs the functioning of the whole organism from a physical point of view and governs all higher cognitive and emotional functions. These are multiple, including perception, attention, memory, executive processes, decision making, language, movement, and sleep. In Tab. 2.1, some functions with their locations are listed. In this thesis, three of these processes have been addressed: movement, sleep, and emotions, which are explained in detail below.

2.2.1 Movements

Movement is controlled by numerous systems with a hierarchical structure that includes CNS and PNS. At the cortical level, the main structures that govern movement are the Primary Motor Cortex (M1), the Premotor Area (PMA), and the Supplementary Motor Area (SMA). They are in close connection with other cortical regions that govern actions and their programming, including hand-eye coordination. In addition, a fundamental structure is the cerebellum that controls fine movements and coordination. An illustrative diagram is shown in Fig. 2.12. The PMA and SMA send information to and receive information from the cerebellum and basal ganglia. Subsequently, these three structures send information to the motor cortex, which in turn sends it to the brainstem, and finally to the spinal cord, which forwards it to the pathways connected to the muscles that will perform the movement. In short, the three cortical structures deal with three connected, complementary but distinct functions. The primary motor cortex is the executor, sending signals to motor neurons to activate muscles. It has a map of the body (motor homonomus, visible in Fig. 2.13) where larger areas are dedicated to parts of the body that require fine movements. The PMA plans and prepares movements in response to external stimuli, selecting the correct action based on the context provided by the other cortices; it also controls the proximal and trunk muscles for posture and orientation. Finally, the SMA coordinates complex movements, especially sequences or

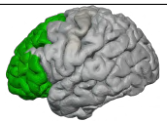
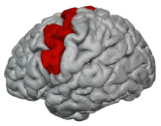
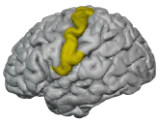
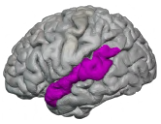
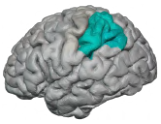
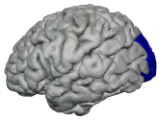
Area	Function	Image
Prefrontal	Higher cognitive functions	
Motor cortex	Control of voluntary movements	
Sensory cortex	Perception and interpretation of primary human sensation	
Auditory cortex	Perception and interpretation of audio stimuli	
Language cortex	Perception and interpretation of language	
Occipital cortex	Perception and interpretation of visual stimuli	

Table 2.1: Structions and functions of cerebral cortex [20].

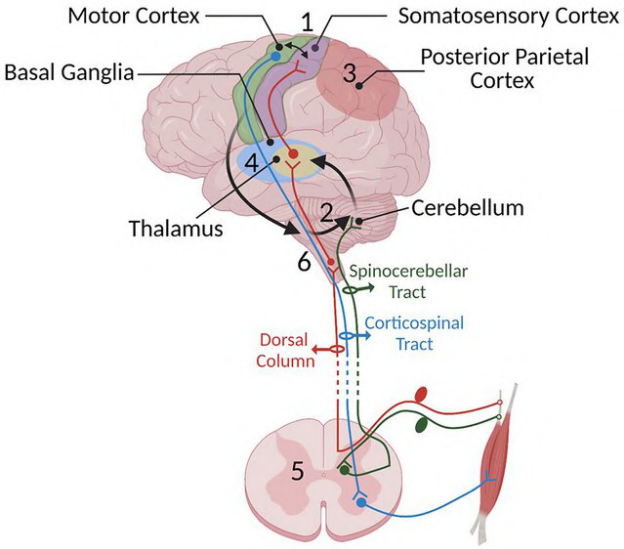


Figure 2.12: Motor pathways. Image from [21].

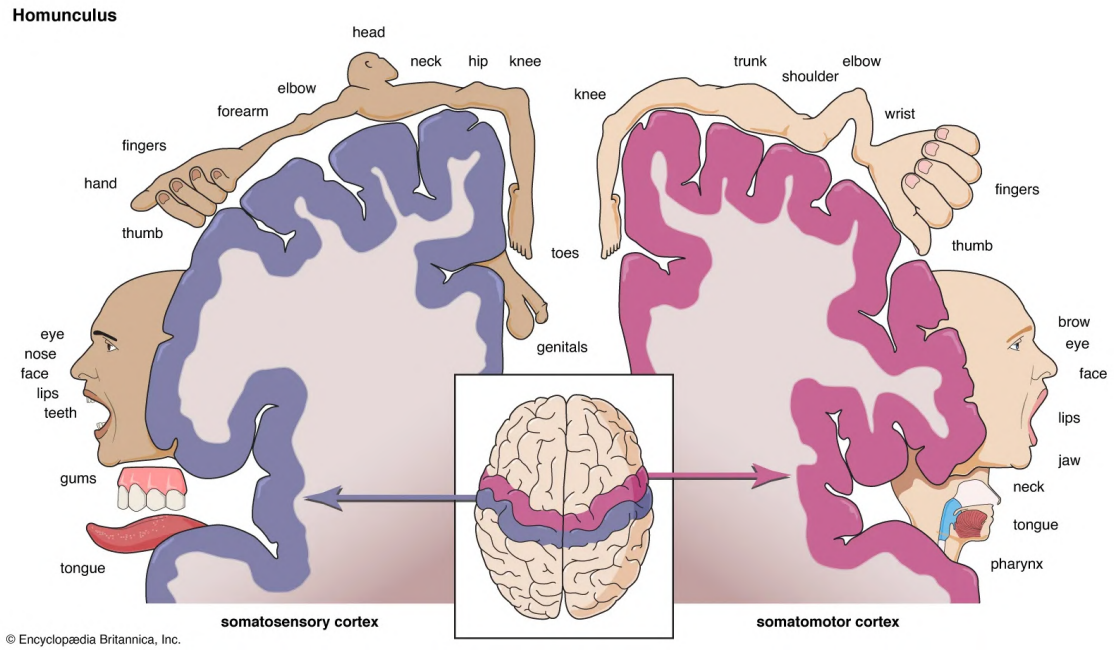


Figure 2.13: The somatomotor and somatosensory homunculus. Image from [22].

actions that require coordination of both limbs. Together, the PMA and SMA form the secondary motor cortex (M2).

From the perspective of EEG signals, the mu rhythm is of particular interest, as it is closely associated with the activity of the motor and PMA. This rhythm, which typically occurs in the frequency range of 8–12 Hz, is characterized by a decrease in amplitude when the subject performs a movement or even prepares to do so. Despite sharing frequency values with the alpha band, this rhythm is known as the mu rhythm because of its well-defined spatial location over the motor and premotor areas. Its suppression has been observed in numerous studies, highlighting its correlation with motor activity and action preparation [23]. During motor activity, a modulation of EEG rhythms is observed, particularly a decrease in the mu rhythm, reflecting the activation of these cortical areas.

2.2.2 Sleep

Sleep is an important physiological stage of daily routine for humans and animals, based on the 24-hour cycle in the presence of an external trigger [24]. In 1968 Rechtschaffen and Kales (RK) [25], based on the analysis of EEG, Electromyography (EMG), and Electrooculography (EOG) parameters, classified sleep into five stages RK classification: four NREM stages (stage 1; stage 2; stage 3; stage 4) and one REM stage. The frequency bands used and the Sleep-related Events in Tab. 2.2 are shown.

The NREM stages are characterized by increasing brainwave synchronization, with the presence

EEG Frequency Bands	Sleep-related Events
Delta Waves (0.5 - 4 Hz)	Associated with the deepest stages of sleep, specifically stage 3 and 4 of NRapid Eye Movement (REM) sleep.
Theta Waves (4 - 8 Hz)	Common during the lighter stages of sleep or the transition from waking to sleeping.
Alpha Waves (8 - 12 Hz)	State of relaxation when the eyes are closed.
Sigma Waves (12 - 15 Hz)	Also known as sleep spindles, these waves are characteristic of stage 2 NREM sleep.
Beta Waves (15 - 30 Hz)	Present during wakefulness but can also be seen during REM sleep.
Gamma Waves (30 - 45 Hz)	Associated with cognitive processes and memory consolidation during sleep.

Table 2.2: EEG Frequency Bands and Sleep-related Events [24].

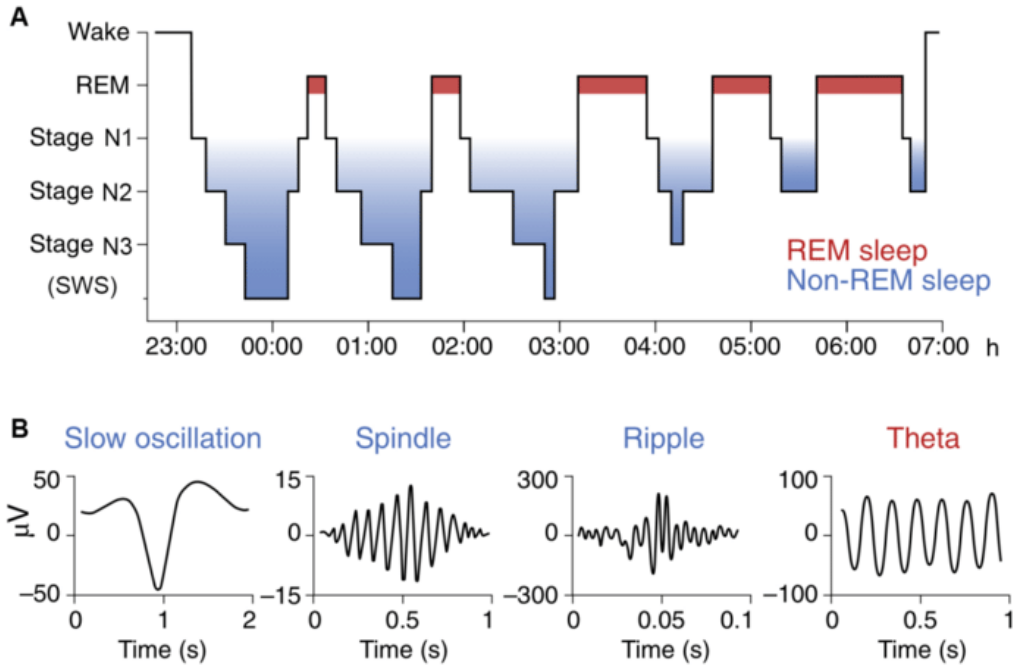


Figure 2.14: The sleep architecture with some elements. Image from [26].

of certain characteristic grapho-elements: NREM stage 1 is similar to wakefulness, NREM stage 2 is characterized by K complex and splinters (shown in Fig. 2.15) and is involved in memory consolidation, NREM stages 3 and 4 are characterized by a strong component of Slow-Wave Activity (SWA) that defines the period Slow-Wave Sleep (SWS). In Fig. 2.14, a normal sleep night is shown with its sleep stages. It is clearly visible that the pattern repeats itself and creates a cyclical pattern, which, however, varies with the passage of the night: initially there is a predominance of SWS that is replaced by the lighter stages of sleep. In Tab. 5.1 a brief description is provided. Compared to the RK classification criteria, the gold standard for the scoring of sleep stages is the

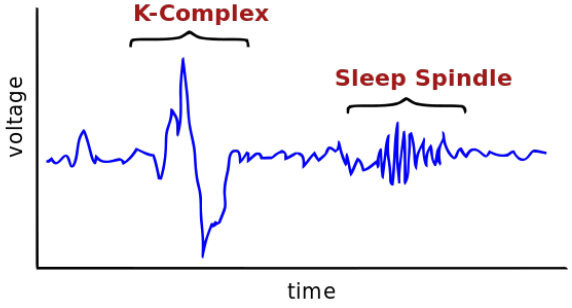


Figure 2.15: Elements of the sleep stage NREM 2. (Author: "Neocadre", Public Domain.)

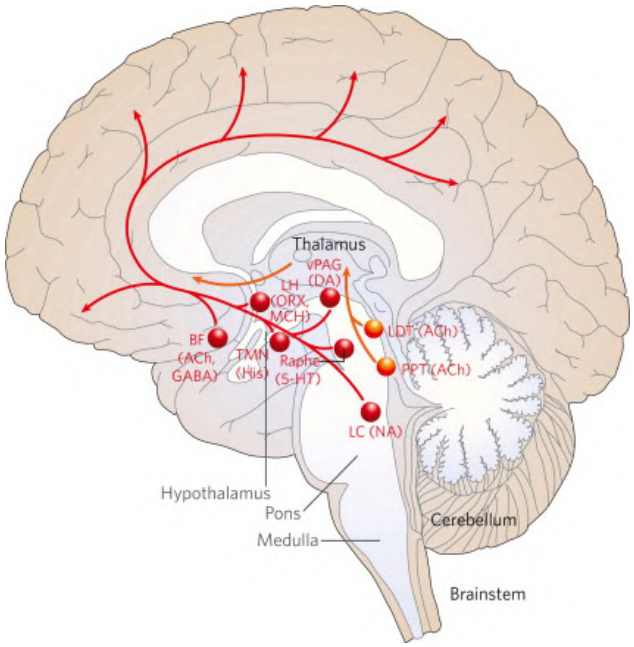


Figure 2.16: The ARAS network. Image from [29].

American Academy of Sleep Medicine (AASM) criteria [27]. This merged stage 3 and 4 of the RK into stage N3, leaving stage 1 as N1, stage 2 as N2 and REM unchanged [28], as shown in Tab. 5.2. The neural circuits of sleep involve a large number of cortical and subcortical structures. One of the main structures is the Ascending Reticular Activating System (ARAS), shown in Fig. 2.16: wakefulness requires the activation of this system, which causes arousal. In REM sleep, visual stimuli (dreams) are generated by stimulation of the reticular formation that sends to the superior colliculus, generating the Ponto-geniculo-occipital waves (PGO), which is useful for marking the onset of REM sleep. Neurons in a portion of the ARAS, send to the thalamus, and a high discharge frequency has been found to determine the waking state or REM sleep; conversely, a low frequency determines NREM sleep. Other nuclei are also involved in the waking and sleeping states, such as the locus ceruleus, raphe, and those of the tuberomammillary nucleus, activated by orexin, a peptide that promotes wakefulness. In addition, a fundamental role is played by the thalamus, which receives much information from the aforementioned structures, and which determines two fundamental states. The first, the oscillatory state, in which neurons synchronize and disconnect from interactions with the outside world (during sleep), and the second, the tonic activity state, is present during wakefulness, where thalamic neurons are receptive to external stimuli. Unstable sleep determines a moderate to severe cognitive deficit according to personal physiology and the duration of the sleep deficit [4, 30]. Some pathologies, determining the deterioration of sleep stages [24] can have a very negative impact on the individual's social, personal, work and physical life [31], and it is necessary to diagnose sleep pathologies in advance to intervene in therapy as soon as possible [24, 32, 33]. Rare pathologies, such as *encephalitis lethargica* [34], lead to the total disappearance of sleep, which inevitably determines the death of the individual after a few weeks [34]. The role of sleep is not well known, although many theories have been developed. One of them, homeostatic theory [35], also known as the "two-process model", explains sleep pressure as the need for sleep to remove metabolic waste products and consolidate synaptic connections built during wakefulness. This model is based on the interaction of two main processes: the sleep process

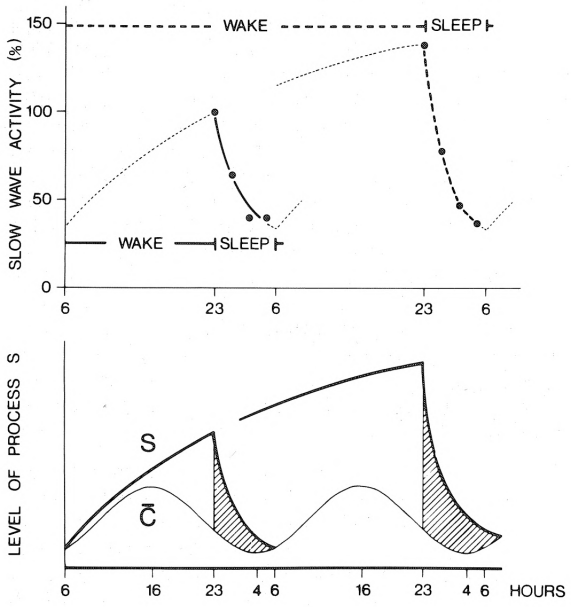


Figure 2.17: The two model sleep stage. Image from [35].

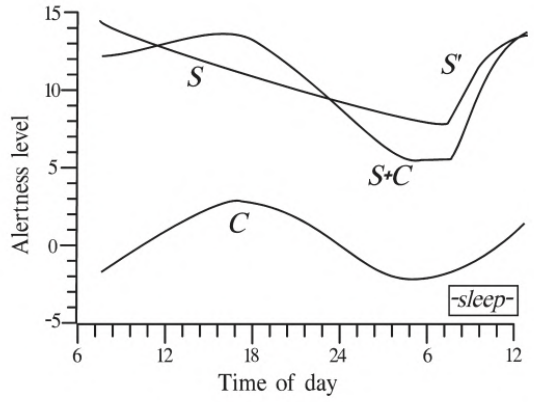


Figure 2.18: The three model sleep stage. Image from [36].

(S) and the circadian process (C). The S process is related to the need for sleep accumulated during the waking period, while the C process is influenced by circadian rhythms, which are regulated by the internal biological clock. This model suggests that the need for sleep progressively increases during the day and decreases during the night, while the circadian process modulates the probability of sleep based on the time of day. The combination of these two processes determines the quality and quantity of sleep that an individual experiences. The diagram in Fig. 2.17 shows the process. An improvement of the Homeostatic Theory is represented by the "three-process model" in which the sleep inertia was included to simulate the daytime alertness [36]. This new model integrates an additional process, the activation process (A), which accounts for the influence of external and internal factors on alertness and attention. This model recognises that in addition to the need for sleep and circadian rhythms, there are also factors that can increase or decrease alertness, such as environmental stimuli, stress and physical activity. The diagram in Fig. 2.18 shows the process.

2.2.3 Emotions

Human emotions have been a subject of interest since the time of Plato, who distinguished intellect, will and emotions. The current idea is that emotions are physiological and cognitive responses to external stimuli or internal mental representations with valence. The characteristics that a process must have in order to be considered an emotion are as follows:

1. Variations occur in at least one response system;
2. Have identifiable triggers that distinguish them from mood;
3. Can be learned or innate;
4. Depend on evaluative processes;
5. Are the result of different neural processes.

The neural systems that govern emotions are multifaceted, starting with the amygdala and in particular the amygdala with both systems, Sympathetic Nervous System (SNS) and PNS. From neuropsychological studies and clinical cases, it is known that the amygdala plays a fundamental role in emotional regulation and that its lesion leads to cases such as that of S.M., who could no longer feel fear. However, Currently no single neural circuit or area has been found that is uniquely and in all subjects devoted to the generation and regulation of emotions, but many circuits and structures are identified as part of the emotional processes. The categorization of emotions has not yet found consensus in the scientific community, and there are currently many different theories that present many different emotional models. The basic theories of emotions can be differentiated into two classes: categorical and dimensional. Categorical ones divide emotions into different classes, while dimensional ones use the assumption that there are dimensions that form a space on which emotions, on the basis of values, are placed and thus there is no discretisation of emotions but they are distributed in a space along a continuum. The earliest theories of emotions were those of James-Lange, who defined emotions as the result of perception and visceral feedback of bodily responses to a stimulus, and that of Cannon Bard, on the other hand, who defined emotions as the result of a cortical stimulus, contemporaneous with the visceral one. Currently, one of the most popular theories is LeDoux's, which presents the hypothesis of two emotion systems in parallel: emotion is generated by neural responses, while feeling is learned through experience. Categorial models, which classify human emotions as discrete entities, are supported by Darwin's evolutionary theory [37]. These models suggest that the most significant emotions have been selectively preserved through evolutionary processes. In subsequent research, Ekman [38] and Plutchik [39] have substantiated this categorical framework by identifying fundamental emotional states, including anger, anticipation, fear, sadness, disgust, trust, surprise, and joy. These foundational states may also be subject to further elaboration [40]. Conversely, dimensional models are theoretically predicated on the interconnections among the dimensions of a conceptual space and the spatial allocation of individual emotions therein, contingent upon the dimensional values under scrutiny. Russel's two-dimensional model [41] is the most widely recognized, as it maps emotions on a two-dimensional plane (Fig. 2.19), with the axes denoting arousal and valence. An ensuing model further incorporates dominance, elevating the structure from a two-dimensional plane to a three-dimensional space, as depicted in Fig. 2.20. Unlike theoretical studies on emotion, experimental

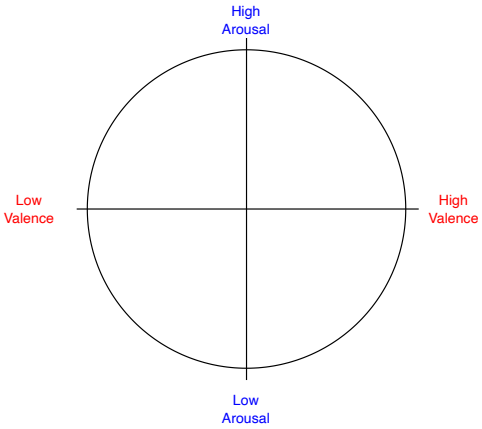


Figure 2.19: The Circumplex Model with Arousal and Valence dimensions

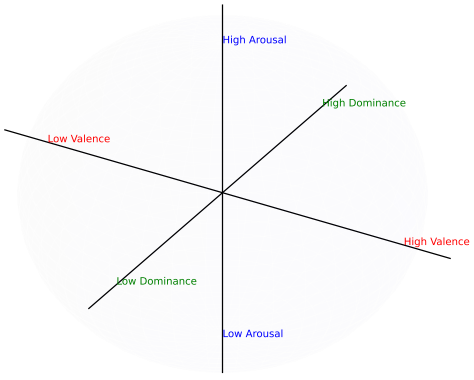


Figure 2.20: 3D Russel's emotional space with Arousal, Valence, and Dominance dimensions (the 2D version of the space contains only Arousal and Valence).

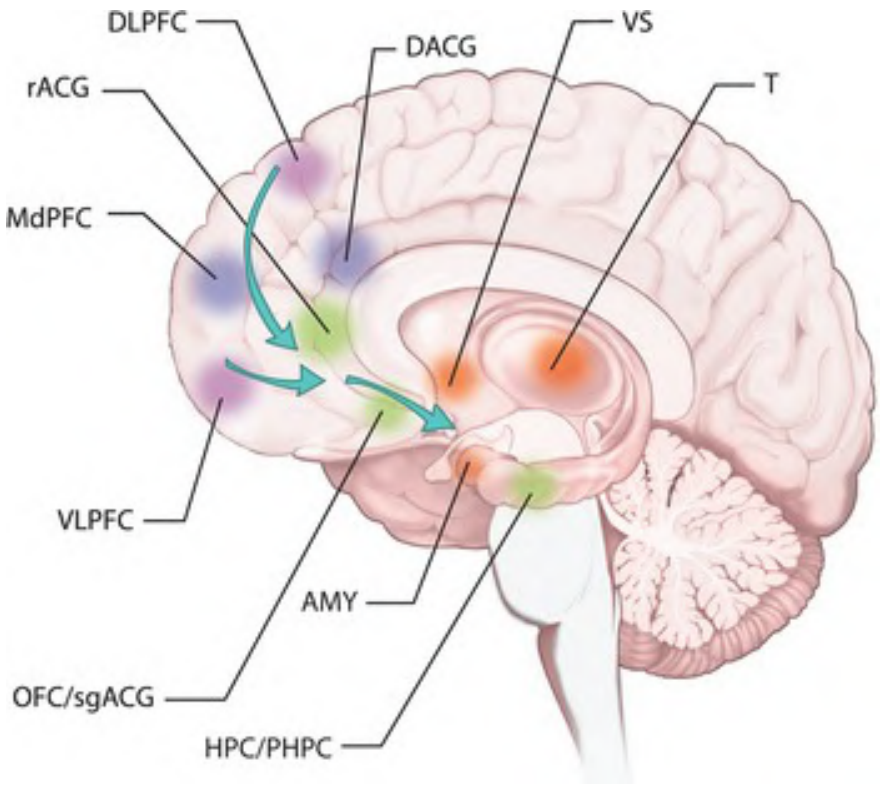


Figure 2.21: The neural circuits involved in emotional process. Image form [42]. As can be seen, the amygdala (AMY) is a small structure that is closely connected to the hippocampus and parahippocampus (HPC/PHPC), which is responsible for memory consolidation processes. This partly explains how memories with a high emotional content are remembered better than others.

neurobiological research supports the fundamental role of the amygdala, visible in the diagram in Fig. 2.21. As can be seen in the figure, the amygdala and the rest of the structures are mainly subcortical, with marginal involvement of the cortical structures. Recent research shows that two neural pathways are involved: the high pathway, through cortical projections, and the low pathway, through subcortical projections. An explanatory diagram is shown in Fig. 2.22. The high pathway, or cortical pathway, involves the processing of emotional stimuli through the sensory cortex before reaching the amygdala. This pathway is characterized by its slower and more detailed processing of information, allowing for conscious awareness and nuanced emotional responses. The connections of the amygdala with the orbitofrontal cortex and other cortical areas facilitate a bidirectional flow of information, allowing the integration of emotional context with cognitive processes [44]. The pathway is particularly important for complex emotional evaluations and decision-making, as it allows the incorporation of contextual and social cues into the emotional response. In contrast, the low pathway, or subcortical pathway, provides a rapid and automatic response to emotional stimuli, particularly those that are threatening or fear-inducing. The subcortical route mainly involves projections from the thalamus to the amygdala, which completely bypasses the cortex. The interaction between these two pathways is vital for a comprehensive understanding of emotional processing. Although the low pathway allows for swift reactions to immediate threats, the high pathway provides a more thorough evaluation of emotional stimuli, allowing individuals to respond appropriately in complex social situations. The amygdala's ability to integrate information from both pathways underscores its central role in emotional processing, as it can respond to coarse and rapid input

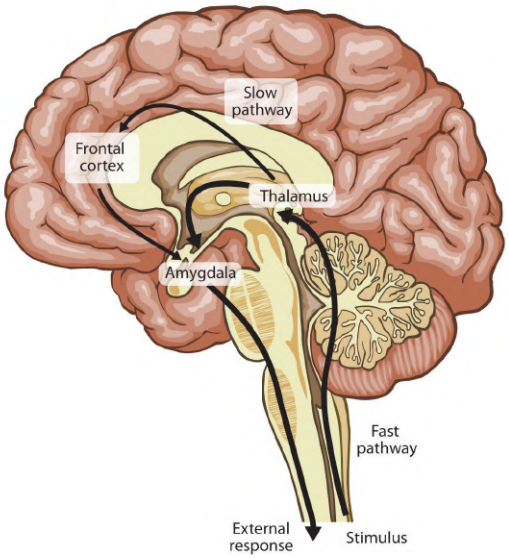


Figure 2.22: Emotional pathways. Image from [43].

from the subcortical pathway while also being influenced by more detailed cortical information [45]. This dual processing mechanism highlights the brain's adaptability in managing emotional responses, balancing speed and accuracy based on the context of the stimuli encountered. Among several biomarkers EEG studied in the context of emotions, Late Positive Potential (LPP) has been observed to be a Event-Related Potential (ERP) component, which will be further illustrated in Section 3.2. It manifests as a sustained positive wave, typically maximal in the centro-parietal areas of the brain, and begins to emerge approximately 300 milliseconds after exposure to an emotional stimulus. This component is particularly interesting due to its association with the processing of emotional and motivational stimuli, showing a greater amplitude in response to emotional stimuli compared to neutral ones [46]. The LPP has been studied in various contexts, revealing that its amplitude is influenced by the valence and arousal of the stimulus. For example, positive and negative emotional stimuli tend to generate a larger LPP compared to neutral stimuli, indicating greater cognitive and emotional engagement [47]. This suggests that LPP not only reflects attention to emotional stimuli but also a deeper and more sustained processing, which can last until the end of the stimulus presentation [46]. In addition, LPP has been used as an indicator of how people regulate their emotions. For example, studies have shown that the use of cognitive re-appraisal strategies can reduce the amplitude of the LPP in response to threatening images, suggesting that people can modulate their emotional reaction through cognitive processes [48]. This makes LPP a useful tool for understanding the dynamics of emotional processing and individual differences in emotion regulation [49]. Finally, LPP has been associated with various psychological disorders, such as panic disorder and anorexia nervosa, where alterations in the LPP response can reflect difficulties in emotional processing and emotion regulation [50]. These observations highlight the importance of LPP as a neurophysiological marker for understanding emotional responses and emotion regulation mechanisms in different clinical populations. Thus, LPP is a crucial ERP component for the analysis of emotional responses, characterized by a higher amplitude in response to emotional stimuli compared to neutral ones, and represents a significant indicator of cognitive processing and emotion regulation.

Chapter 3

Brain Activity Measurement and Analysis

In this chapter, a brief introduction to the methods for investigating the activity of CNS and the field of BCI is provided, including techniques, algorithms, and applications, with a greater focus on EEG, a technique used throughout the work presented in this thesis. The information in this chapter is taken from [4–10, 51–60].

3.1 Neuroimaging techniques as a whole

There are various techniques for studying CNS and each one has different advantages and disadvantages. They can be divided into two macrocategories: those that use the electricity produced by CNS, and those that use metabolic and blood information. In the first category, instruments such as EEG, Intracranial Electroencephalography (iEEG), and Electrocorticography (ECoG) measure the electrical activity of CNS. Although they offer exceptional temporal resolution, their spatial resolution is limited. Furthermore, the dispersion of electrical currents poses challenges in pinpointing the exact location of the source [61]. In the second category, fMRI and functional Near-infrared Spectroscopy (fNIRS) assess metabolic activity through different physical principles. Specifically, fMRI exhibits superior spatial resolution accompanied by very poor temporal resolution and represents a significantly costly apparatus. In contrast, fNIRS is more economical than fMRI, but its penetration capacity is restricted to just a few millimeters. In this work, the EEG device was used, which will be explained in more detail at the end of this paragraph.

Continuous monitoring of the neural signal can be very useful in the case of rehabilitation after surgery or in the case of BCI permaments for severely disabled people. The ECoG technique [62] is invasive and allows the signal from the cerebral cortex to be recorded by electrodes applied under the scalp, as can be seen in Fig. 3.1. In addition, this technique offers a higher spatial resolution than classical EEG, being directly in contact with the brain surface. Compared to EEG, ECoG has a wider registration band and much less noise than EEG. A classic use of ECoG is the detection of epileptogenic outbreaks.

Among invasive techniques, it is iEEG (also called StereoEEG), which is the recording of the

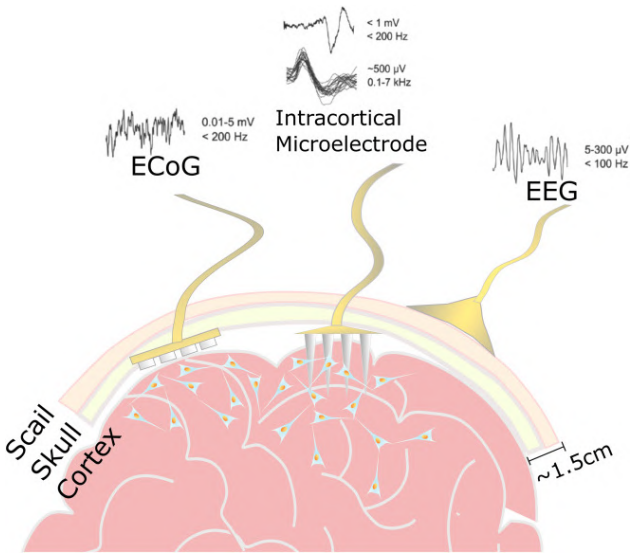


Figure 3.1: The three different ways to record electrical activity from the brain. Image from [62].

electrical activity of a precise sector of the brain using intracerebral electrodes. This technique, which is often used in subjects with epilepsy [63], can be useful in identifying epileptic foci and combined with clinical tools can calm epileptic seizures and prevent them through surgery. It has a high spatial resolution because it is positioned exactly in the area of interest.

Magnetoencephalography (MEG) is a neurophysiological recording technique that measures the magnetic fields generated by neuronal electrical activity in the brain: the correte generated by the sum of post-synaptic potentials generates a magnetic field [64]. This method offers numerous advantages over EEG. One of the main strengths of MEG is its high temporal resolution, which allows brain events to be recorded with millisecond-by-millisecond precision, making it particularly useful for studies of rapid brain dynamics and for real-time localization of brain activity [65]. In addition, MEG has a higher spatial resolution than EEG, as magnetic fields undergo less distortion when passing through biological tissues, allowing better localization of brain sources [65]. This is crucial in clinical applications, such as the location of epileptic foci or mapping of brain function prior to surgery [66]. The operation of MEG is similar to that of EEG: It is possible to extract Evoked Potential (EP), and also Entropy Rate Minimization (ERM) (the magnetic equivalent of ERP), with the advantage that it can be more precise in terms of localization, since it does not suffer as with EEG from the current distribution along the scalp surface, and a reference electrode is not required. A major limitation is the high cost and complexity of the required equipment, which requires controlled environments to reduce external magnetic noise. Furthermore, while MEG is effective in detecting superficial brain activity, it can be difficult to accurately locate deeper sources of neuronal activity, such as those located in the brainstem or subcortical structures [67]. Another critical aspect is the need for advanced data analysis techniques to extract meaningful information from MEG signals, which may require specialized skills and time [65]. In comparison, EEG is a more accessible and less expensive technique that is widely used due to its ease of use and the ability to record brain activity in a variety of clinical and research contexts. However, EEG has a lower spatial resolution than MEG, as electrical signals undergo significant distortions as they pass through the skull and brain tissues, which can lead to inaccurate localization of brain sources. Furthermore, EEG is subject to artifacts from eye and muscle movements, which can further complicate the

interpretation of the data [68]. An example of the instrumentation is shown in Fig. 3.2, and in Fig. 3.3 the operation of MEG is shown.



Figure 3.2: The MEG instrument. (Author: "NIMH author", Public Domain).

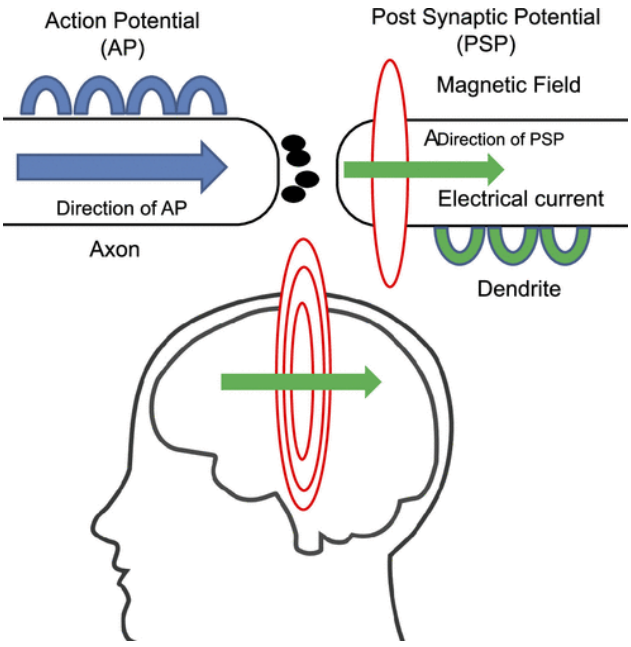


Figure 3.3: A brief description of how MEG works. Image from [64].

A non-invasive, yet commonly used technique in the analysis of the CNS includes Magnetic Resonance Imaging (MRI) and fMRI. The former is useful for studying the anatomy of the CNS while the latter is used for functional studies via the Blood-Oxygen-Level-Dependent (BOLD) signal. Unlike EEG, fMRI has excellent spatial resolution yet poor temporal resolution, which is why these two techniques are often used together. MRI works by using strong magnetic fields and radio waves to generate detailed images of the organs and tissues in the body. The MRI device consists of different components that contribute to generating the image: (1) the main magnet generates a powerful magnetic field that aligns the nuclei of hydrogen atoms in the body; (2) the gradient coils produce magnetic linearly varying in space and oriented along the three spatial axes; (3) the shim coils correct field inhomogeneities; (4) radiofrequency pulses send short bursts of Radiofrequency (RF) energy towards the aligned nuclei, disturbing the alignment. Imaging starts by producing a very strong magnetic field, typically from 0.5 T to 1.5 T for clinical applications. The latest scanners can generate, for research purposes, a static magnetic field as high as 7.5 T. Gradient coils generate a supplementary magnetic field which changes linearly in space so that every point of the area to be imaged will experience a different magnetic field. First of all, an strong static magnetic field is generated (B_0), which aligns the hydrogen nuclei (^1H, protons) normally oriented at random. Protons within the water molecule are exploited since water is among the most plentiful components within the human body though the distribution and interaction varies from one tissue to the other. Protons possess magnetic properties basic to the creation of an image. The (B_0) field aligns the protons in the direction of the field and thus causes longitudinal magnetization. Then, a RF pulse of the same precession frequency as the protons produces another magnetic field, (B_1), that changes the orientation of the protons from their aligned position for a short period of time. While returning to their position, they emit a signal called Free Induction Decay (FID). The amplitude and duration of the FID are dependent on two factors: T1 relaxation time (T1) relaxation time, or T1, and T2 relaxation time (T2) relaxation time, or T2. The T1, or spin-lattice relaxation, is the time taken by the protons to release 63% of their longitudinal magnetization and realign themselves with the (B_1) field. The T2, or spin-spin relaxation, is the time taken by the protons to lose 37% of their transverse magnetization and resynchronize with their initial phase. The decay of transverse magnetization depends on such factors as interactions between protons and surrounding tissues and inhomogeneities in the (B_0) field. These phenomena form the basis for obtaining detailed and tissue-specific diagnostic images. fMRI is based on MRI principles and technique, but instead of visualizing the structure of the brain, it detects activity in the brain by acquiring the signal of blood flow and oxygen supply. The resultant signal is called the BOLD signal. The reason for this is that blood is paramagnetic, its properties depending upon whether it is oxygenated or deoxygenated. Haemoglobin is an iron-carrying protein in the blood, and when it binds oxygen or fails to do so, its magnetic nature changes. The two forms of haemoglobin include oxygenated (oxyhaemoglobin) and deoxygenated haemoglobin; the former is diamagnetic while the latter is paramagnetic. This is the fMRI technique to analyze which brain region requests more oxygen and points to the area being more active during some tasks. In Fig. 3.4 the differences between the MRI and fMRI images is shown.

fNIRS is a non-invasive way to view the happenings of the brain. It involves the use of near-infrared light (650-1000 nm), a type of light invisible to human sight, which can pass through skin and skull. It has a light source and a detector [70, 71] Similar to fMRI, fNIRS depends on monitoring

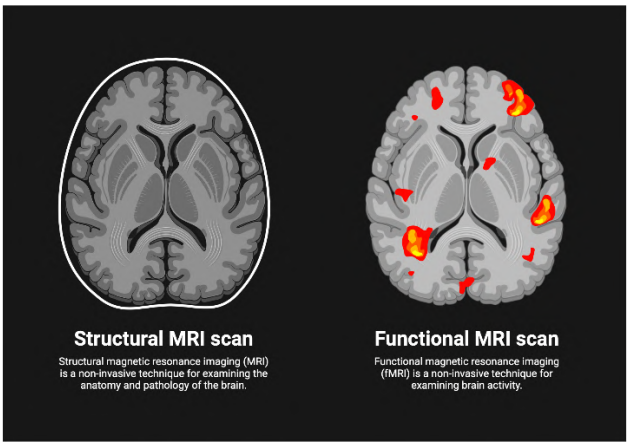


Figure 3.4: (left) A structural MRI and (right) an fMRI acquistion on the MRI after the skull removing. Image from [69]



Figure 3.5: An EEG-fNIRS instrument. Image from BrainProducts®website.

the flow of blood within the head. The harder a region of the brain is working, the more oxygen it needs. Thus, by monitoring where oxygen-rich blood is flowing it is possible to ascertain which parts of the brain are active. fNIRS is able to detect the differences between oxygenated (O_2Hb) and deoxygenated (HHb) blood. This becomes important in describing the utilization of oxygen by the brain at any one time. fNIRS is great, in that it gives us a good picture of where things are happening in the brain. However, it's not so good at tracking very fast changes, since it takes a little while for the blood to flow to the active areas. The Fig. 3.5 shows an fNIRS device being used together with an EEG, which is another way to study brain activity.

In principle, the above measurement techniques could be all used for collecting the brain activity usable in BCI, though most of them are expensive, cumbersome, difficult to wear, and usable only in laboratory environments. EEG, among them, is a light and low-cost technique that can be easily worn and used in real-life situations and, for this reasons, is the one used in this thesis and deserves a specific description.

3.2 Electroencephalography (EEG)

The EEG is a non-invasive technique that allows recording of the brain's electrical signal using electrodes placed over the scalp. The technique was discovered in 1875 by Richard Caton who

recorded the brain waves of animals and was then used for the first time by Hans Berger in 1929 on humans, recording the first EEG in history, followed by the work of Moruzzi and Magoun on sleep-wake cycles. It is a technique that records cortical electrical activity primarily generated by cortical layers III, V, and VI, and it is difficult to record subcortical activity by surface EEG without the integration of subcortical techniques or the help of other techniques such as fMRI [72]. The measured currents are the result of the sum of the post-synaptic potentials of the population discharging synchronously, and this makes it possible to measure the electric potential on the scalp that would otherwise be too low to be measured. Numerous recording systems are subdivided according to different characteristics. The first subdivision can be made according to the use of electrodes, which can be either *active* or *passive*: the former is the *active* electrodes that are amplified, or *passive*, i.e. without amplification. A second subdivision is according to whether the gel is used or not, *wet* or *dry*. To obtain a signal of good quality, the impedance of the electrodes should be between 1Ω and 10Ω: dry electrodes facilitate the recording but the progressive drying during the recording, and thus an impedance gradient; *wet* electrodes require more care during assembly, do not allow impedance levels comparable to the former and are more susceptible to individual variations, but do not suffer from the problem of gel drying and loss of conductivity. EEG has two main configurations, unipolar and bipolar, depending on the potential difference to be measured. The 10-20 standardized system is commonly used for electrode placement [73], as shown in Fig. 3.7. The standard ensures a comparison between different data acquired in different centers and with different instruments, although there is a good deal of individual variability between individuals that does not allow a perfect overlap of areas and electrodes. In Fig. 3.6 an example of EEG data from the public dataset is shown.

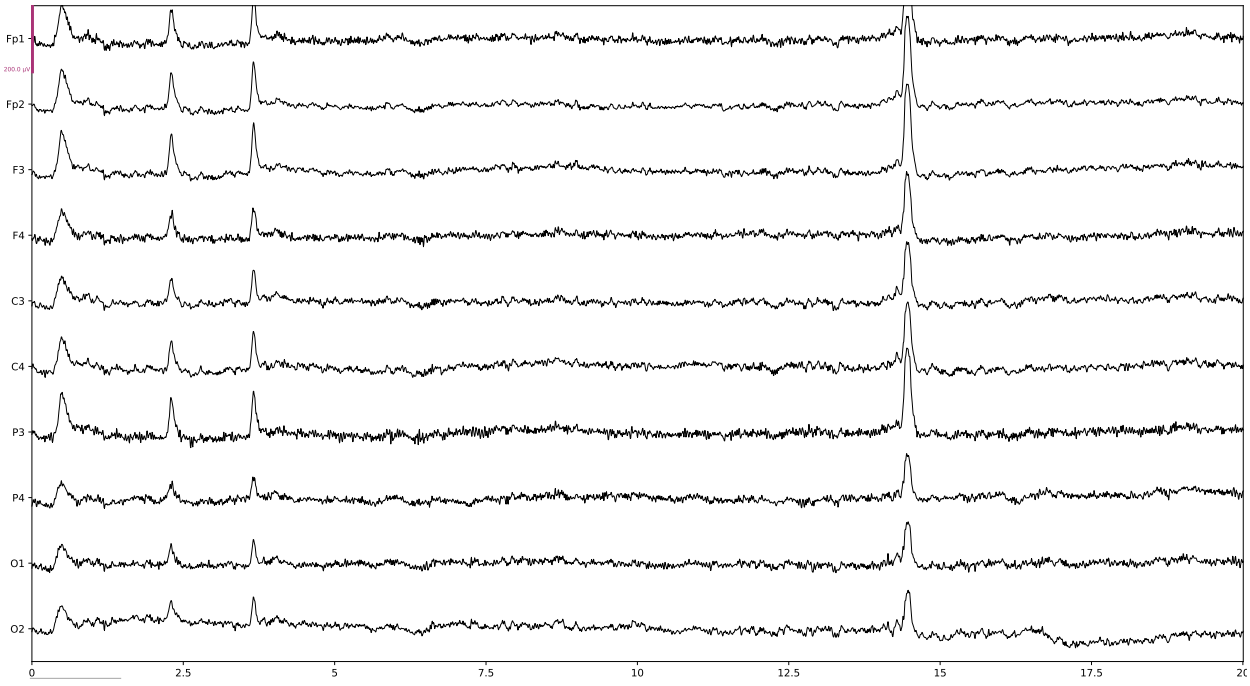


Figure 3.6: EEG data. On y-axis, for each row (channels) the amplitude in μV, on x-axis, the time in seconds. The spikes are eye movement artifacts.

The study of the signal EEG is done in terms of frequency and amplitude. The rhythms of EEG studied mainly are as follows:

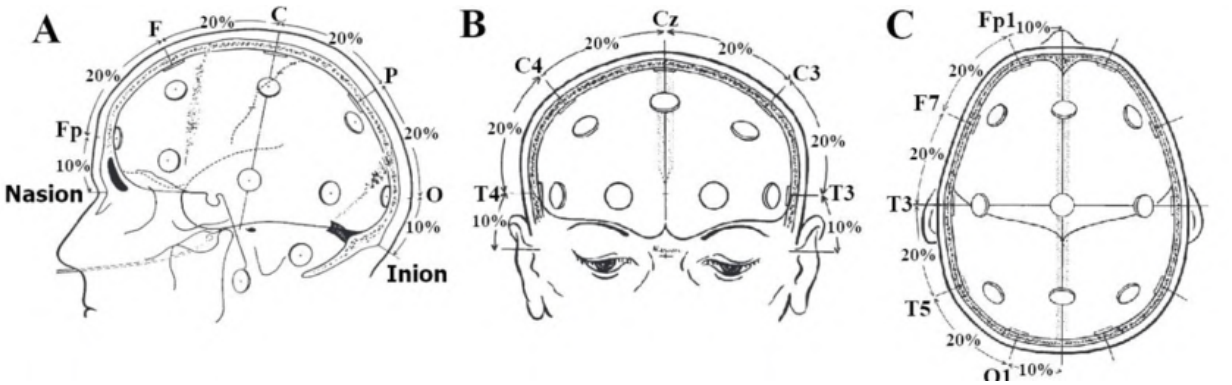


Figure 3.7: Standardized position of electrodes. Image from [74].

- **Delta waves (< 4 Hz)**: Associated with deep sleep, anesthesia, and coma states.
- **Theta waves (4 - 8 Hz)**: Present during light sleep, meditation, and states of deep relaxation.
- **Alpha waves (8 - 12 Hz)**: Dominant during the relaxed waking state, with eyes closed. They indicate a state of calm and mental rest.
- **Mu waves (8 - 12 Hz)**: Similar to alpha waves, but located in the motor areas of the brain. They are present at rest and decrease during movement or observation of movement.
- **Sigma waves (12 - 15 Hz)**: Present mainly during Non-Rapid Eye Movement (NREM) sleep, particularly in phase 2, and are associated with memory consolidation [75].
- **Beta waves (12 - 30 Hz)**: Associated with active mental activity, concentration, problem solving, and decision making. They are subdivided into:
  - **Low beta (12 - 15 Hz)**: Associated with states of relaxed attention and internal focus.
  - **High beta (15 - 30 Hz)**: Associated with states of high concentration, alertness, and even stress and anxiety if excessive.
- **Gamma waves (> 30 Hz)**: Involved in complex cognitive functions such as learning, memory, and information processing. They are subdivided into:
  - **Low range (30 - 60 Hz)**: Associated with sensory processing and information integration.
  - **High range (> 60 Hz)**: Associated with higher cognitive processes, such as consciousness and self-awareness.

Many information can be extracted from the data. Conventionally, the background rhythm, mean amplitude, regularity, symmetry, position, morphology, and pattern of the graph-elements are analyzed. Among the most classic analyzes that can be performed on the signal EEG, EP are bioelectric events that can be recorded following events. They are divided into two categories: Event-Stimulus Potential (ESP) and ERP. The former are events that can be detected on the trace that depend exclusively on the characteristics of the physical stimulus (amplitude, localization, latency): they can be visual, sensory, and acoustic. The Steady-State Visually Evoked Potential

(SSVEP) [76] constitute a peculiar class of ESP, representing typical responses to visual stimuli at defined frequencies. When the retina is exposed to visual stimuli between 3.5 and 7.5 Hz, the brain produces electrical signals that correspond to the same frequencies (or multiples thereof). The ERP, on the other hand, are events that depend on the information within the stimulus, such as the P300 Event-Related Potential (P300), the Contingent Negative Variation (CNV) and the Motor-Related Cortical Potential (MRCP). The P300 is elicited by rare and significant stimuli that require cognitive evaluation or decision making. It is associated with processes of attention, stimulus discrimination, and memory updating, and is mainly recorded in parietal regions. The CNV is a slow, negative response that develops during anticipation of an expected event, usually an imperative signal that requires a motor response. It reflects expectation, motor preparation, and attentional control and is mainly recorded in the central and frontal regions. The MRCP is a positive deflection that occurs at the onset of movement. It is related to the activation of cortical motor areas and the execution of the movement, and is recorded mainly on the motor cortical areas. Finally, among the main ones is Mismatch Negativity (MMN), a negative deflection that appears when a rare deviant stimulus is present compared to the average stimulus. An example of P300 is shown in Fig. 3.8. The

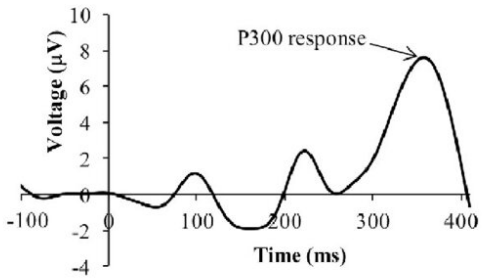


Figure 3.8: P300 wave. Image from [77].

amplitude of the EP is only a few μV compared to the fluctuation of the background signal, which is composed of a stochastic signal of tens of μV, so to extract the graph-element it is necessary to eliminate the background rhythm by averaging many trials in order to bring out the constant element. The recording of several sessions thus serves to reduce uncorrelated noise. This process can be formalized by Eq. 3.1:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x_i^c(t) = 0 \tag{3.1}$$

where:

- $x_i^c(t)$ represent the i time-series of EEG signal recorded by the channel c as a function of time t,
- N is the total number of time series included in the average,
- the average is calculated *channel-wise*, i.e. separately for each channel c of the EEG signal,
- the result tends to zero (0) when N tends to infinity, assuming that the signals are centred and that the noise is uncorrelated between the time series.

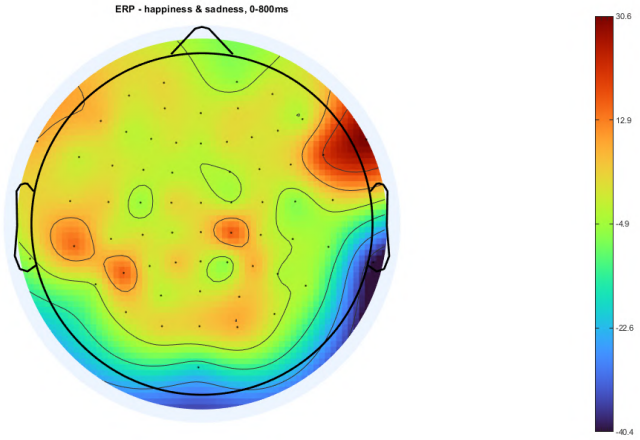


Figure 3.9: Example of Topo-ERP from Chapter 6.

This phenomenon reflects the fact that, by increasing the number of time series considered, averaging tends to eliminate random noise components, preserving only the common or coherent signal between channels. In practice, this property is exploited to improve the Signal-to-Noise Ratio (SNR) in EEG signals, which is generally low due to its nature (artefacts, instrument instability and low amplitude). The applications of ERP are multiple: they can be used to study attention processes and the amplitude of ERP can be used as an indicator of the level of attention. In addition, they can be used to analyze stored memory traces, and P300 is widely used in this area. Finally, for the study of language, N400 Event-Related Potential (N400), a negative deflection wave is used in the study of the semantics of words within a sentence, and its amplitude varies depending on its congruence with speech. The topographical representation of ERP is called Topographical Event-Related Potential (Topo-ERP), and an example is shown in Fig. 3.9.

3.2.1 Physiological Artifacts

Artifacts are signal fluctuations that, in the EEG domain, can be due to two main categories: *physiological artifacts* and *external artifacts*. The removal of artifacts is one of the fundamental pre-processing steps in the data, and improves the accuracy of classification by an automatic classification and prediction system.

Ocular: artifacts due to ocular displacement are among the most common [78]. An average person blinks between 14 and 17 times per minute [79], and furthermore, the movement of the eyes causes a change in the electrical potential because the potential difference between the retina and the cornea creates a dipole that influences the signal EEG due to the volume conduction effect, and furthermore, the amplitude of a signal generated by eye movement is many times greater than that of EEG. [79].

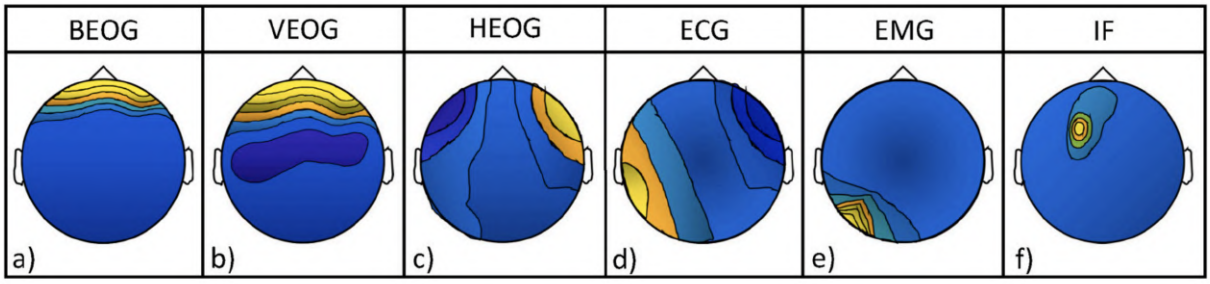


Figure 3.10: Topographical maps constructed on the basis of Independent Components (ICs) calculated by ICA. Image from [83]. BEOG: blink (a); VEOG: vertical eye movement (b); HEOG: horizontal eye movement (c); ECG: cardiac artifact (d); EMG: muscular artifact (e); IF: high impedance (f).

Muscular: muscle contraction causes many artifacts and is distributed in the 0 - 200 Hz band, and the amplitude, shape, and duration depend on the contraction and the location of the muscle [80].

Cardiac: the heart is an involuntary striated muscle that uses an electrical impulse to contract. The first cardiac artifact is the one due to the pulsation of the blood vessels, at ~ 1.2 Hz, and whose waveform is very similar to that of EEG; the second artifact is given by the electrical signal of the heart, which has a regular shape and can be removed more easily [81, 82]. In Fig. 3.10 some artifacts extracted using Independent Component Analysis (ICA) and using the topographical representation of the ICA unmixing matrix are shown.

3.2.2 External Artifacts

External artifacts may be due to other instrumentation, current in city power lines, instrumentation faults, or the recording instrument of EEG itself. An inherent difficulty of EEG is to differentiate an activity between two areas (volume coherence) and an artifact that propagates in different areas.

3.3 Computational tools for EEG

3.3.1 Signal decomposition/transformation

There are numerous signal decomposition techniques that can be used in the removal of signal artifacts EEG, some are more precise but slow while others are more oriented towards online artifact removal systems. One of the main methods of artifact removal is the use of bandpass and notch filters. The traditional regression method assumes that each channel has a portion of pure signal EEG and a certain amount of signal due to artifacts: first, the relationship between the amplitude of the reference channel and the channel EEG is defined using a transmission factor [80, 84], and then the estimated artifact is subtracted. For the application of this method, it is necessary to use external channels of the artifacts to be removed. The method is shown in Eq. 3.2.

$$EEGcor = EEGraw - \gamma Factor(ArtifactChannel) \tag{3.2}$$

where $EEGcor$ is the EEG signal, $EEGraw$ is the raw EEG signal, $\gamma Factor$ is the transmission factor between the two channels, and $ArtifactChannel$ is the reference channel of the artifact to be eliminated. The use of Wavelet Transform (WT) is also a widely used method and compared to the use of Fourier Transform (FT) has the advantage of having a better compromise in the display of time-frequency information for non-stationary data, such as EEG. Wavelets are waveforms with a time component, zero mean value, and nonzero energy. In WT the signal is decomposed using a series of functions that are variations of the mother wavelet and allow different information adapted to different frequencies to be extracted. There are many families of wavelets, and the most widely used in EEG is the "Morlet". The Discrete Wavelet Transform (DWT) is defined by the equation 3.3:

$$\psi_{j,k}(t) = \frac{1}{\sqrt{a_0^j}} \psi \left(\frac{t - k b_0 a_0}{a_0^j} \right) \tag{3.3}$$

where j and $k \in \mathbb{Z}$, $a_0 > 1$ is a fixed dilation step and b_0 is a translation factor that depends on a_0. The set of techniques Blind Source Separation (BSS), a series of unsupervised algorithms, is grouped because of the idea that information can be separated without appropriate information from external channels. The general idea is that signal X is the observed signal, signal S is the source signal including useful signal and artifacts, and A is a matrix that linearly mixes the signals (Eq. 3.4)

$$X = AS \tag{3.4}$$

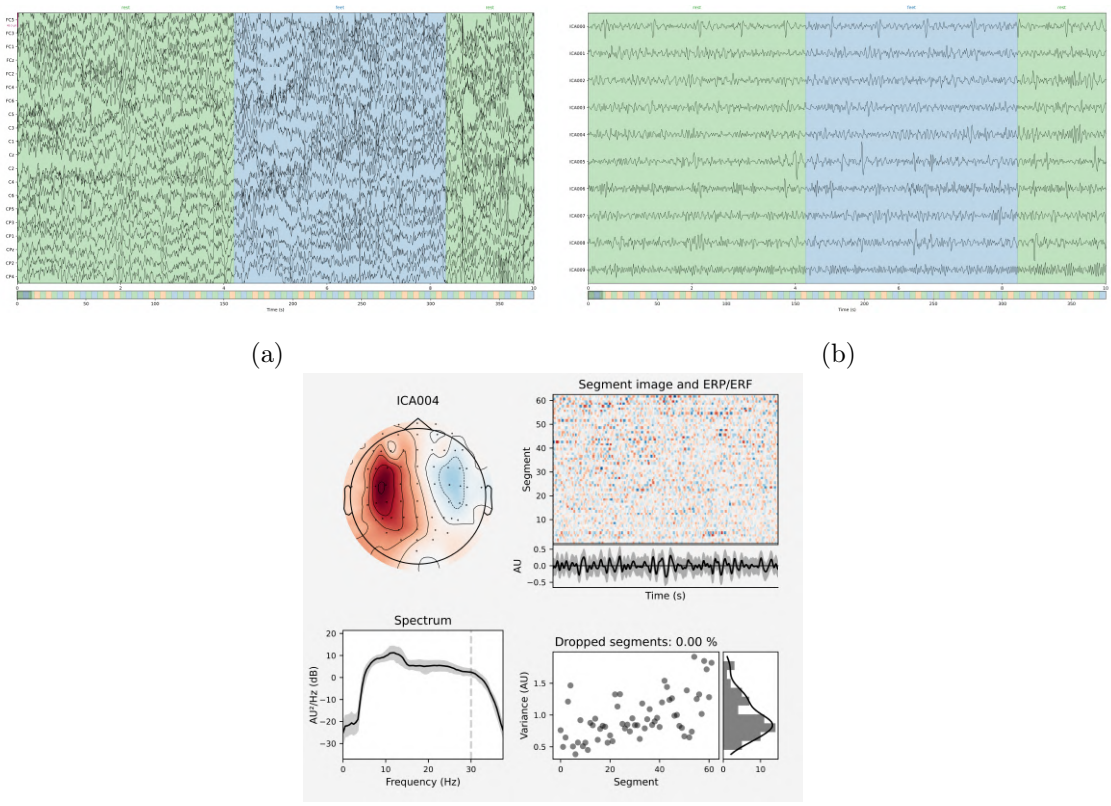
These signals can be separated by finding a matrix W that multiplied by X, returns U, the estimate of the original sources, which can then be removed to reconstruct the original signal (Eq. 3.5):

$$U = WX \tag{3.5}$$

Among these techniques, there is the Principal Component Analysis (PCA), which is based on the eigenvalues of the covariance matrix and allows the conversion of correlated variables into uncorrelated variables called Principal Components (PCs). Another widely used technique in the BSS family is the ICA technique, which assumes that signal sources EEG are a linear mixture of brain signal artifacts and can be divided into ICs. The basic assumptions of this technique are as follows:

- The signals are statistically independent of each other;
- The observed signals are equal to or greater in number than the sources to be extracted;
- The sources do not have a Gaussian distribution or at most one source may have a bell-shaped distribution.

There are numerous algorithms for the extraction of ICs, (FastICA, Picard, etc.) and many algorithms based on AI for the removal of artifacts based on ICA, including the IC-Label [85] or the method based on the topographic representation of ICs [83]. An example, using a public dataset [86], of the original EEG signal and the relative ICs extracted by ICA is shown in Fig. 3.11. Moreover, the topographical component of one IC is also shown as an example and its relative properties.



(a) (b)

(c)

Figure 3.11: Example of ICA decomposition results: (a) raw EEG signal, (b) ICA sources, (c) example of one IC with its properties. The different background colours in (a) and (b) represents different mental tasks (green: rest; blue: feet movement)

Another widely used technique is Canonical Correlation Analysis (CCA), which separates components from uncorrelated sources by finding linear relationships between two variables by maximizing the pairwise correlation between two time series. In addition, Empirical Mode Decomposition (EMD) [87] is a technique used to decompose the signal EEG, and is a heuristic technique specifically designed for nonlinear and nonstationary signals [88]. It is a data-driven technique that decomposes the signal via an iterative algorithm into defined components Intrinsic Mode Functions (IMFs), using local maxima and minima and subtracting from the mean until a constant residual signal or monotonic function is obtained. The EMD works in the following way: Given *s(t)* as signal:
(1) extract local maxima and minima from *s(t)*, (2) compute the envelope ϵ(t) for maxima and for minima, (3) compute the average of two envelope α:

$$\alpha(t) = \frac{\epsilon_{max}(t) + \epsilon_{min}(t)}{2} \tag{3.6}$$

Extract *d(t)* by:

$$d(t) = s(t) - \alpha(t) \tag{3.7}$$

Check if *d(t)* satisfy the condition for IMF Repeat the iteration. The new signal is:

$$g(t) = s(t) - IMF_1(t) \tag{3.8}$$

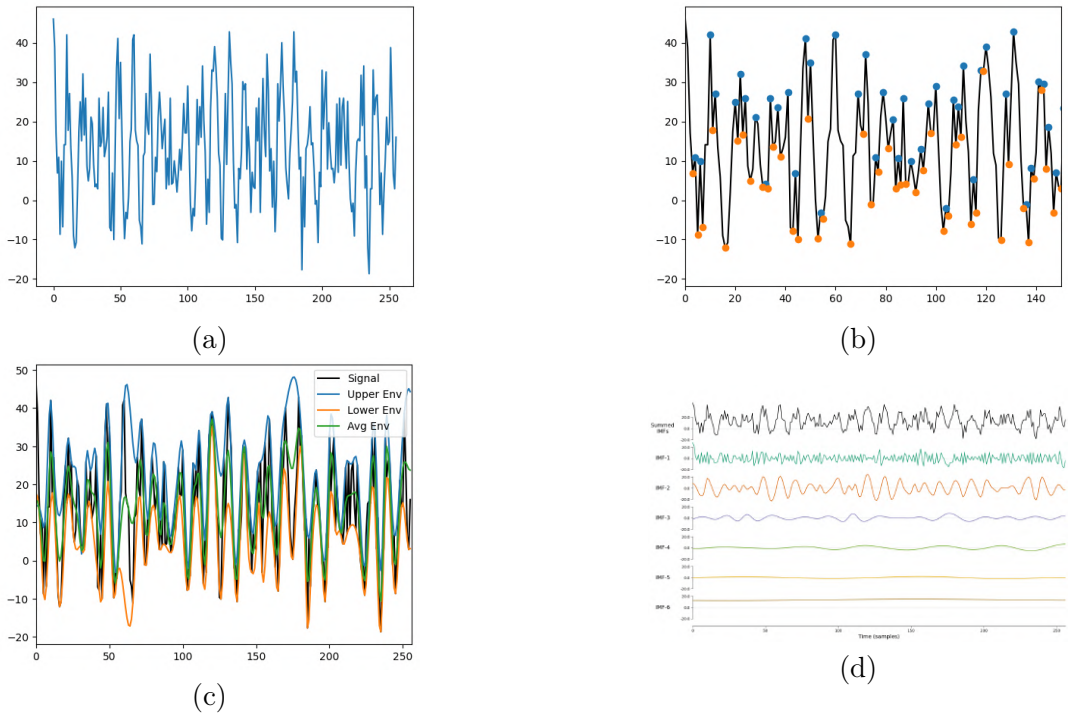
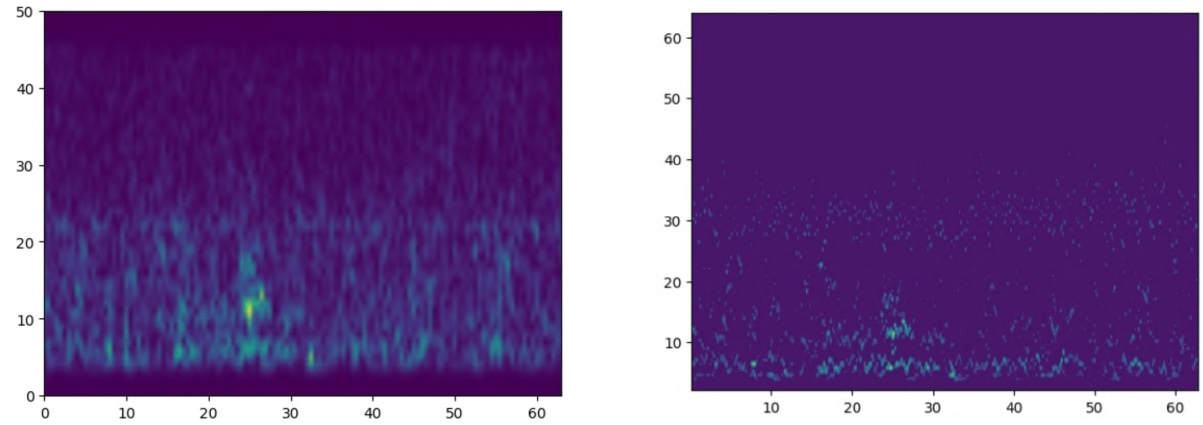


Figure 3.12: An example of the EMD decomposition: (a) original EEG signal, (b) local maxima and minima, (c) average computing, (d) the process is repeated until the function obtained is a monotonic function, and the results are the IMFs.

The process is repeated until the function obtained is a monotonic function and the new signal is:

$$s(t) = \sum_{i=1}^{M} IMF_i(t) + R(t) \tag{3.9}$$

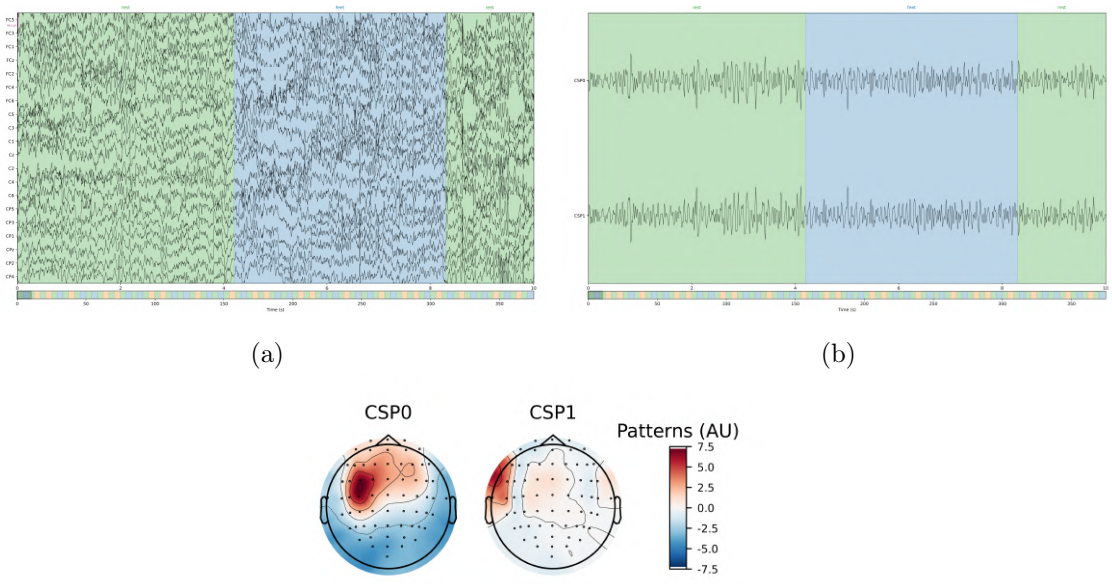
where M is the number of IMFs and $R(t)$ is the residual of the signal after decomposition. An example of the EMD decomoposition is shown in Fig. 3.12 using PyEMD [89].
There are therefore numerous variations of EMD, including Ensemble Empirical Mode Decomposition (EEMD) [90], Multivariate Empirical Mode Decomposition (MEMD) [91], Fast Iterative Filtering (FIF) [92], with its extension Multivariate Fast Iterative Filtering (MvFIF) [93, 94], its time-frequency representation called IMFogram [95] and Adaptive Local Iterative Filtering (ALIF) [96]. A representation of the EEG signal from Dataset for Emotion Analysis using Physiological signals (DEAP) dataset [97] using IMFogram and using time-frequency representation with FT is shown in Fig. 3.13.



(a) (b)

Figure 3.13: Time-frequency representations using **(a)** FT and **(b)** MvFIF showed though IMFogram of EEG data from DEAP dataset [97].

A widely used technique in the study of signal EEG, which is mainly applied in the study of motion, is Common Spatial Pattern (CSP) [98]. A statistical technique used to analyze multichannel data, often employed in brain-computer interfaces to extract features from EEG signals. Objects are identified as spatial patterns that maximize variance between two conditions or classes, such as different mental states or motor imagination tasks. Calculate spatial filters that project the data onto a low-dimensional subspace where differences between classes are most evident. In Fig. 3.14, an example of the components of CSP extracted by an EEG signal is shown.



(a) (b)

(c)

Figure 3.14: Example of CSP decomposition results: (a) raw EEG signal, (b) CSP components and (c) CSP patterns. The different background colors represent different mental tasks (green: rest; blue: feet movement)

3.3.2 Signal processing

The study of the signal EEG can be done in two different ways: by studying individual channels and transforming the signal channel by channel (channel-wise), or by studying connectivity. In the first case, one can study different signal features, such as mean, variance, peaks, and extract ERP. Furthermore, signal transformations from the space domain to the frequency domain can be used using mathematical techniques such as FT or WT. The FT is a fundamental tool in signal analysis, and since analog signals are converted into digital signals for processing, the term FT will henceforth be used to refer to the discrete FT (also known as the Discrete Fourier Transform in the digital domain). Given a sequence of length N, $x[n]$ with respect to $n = 0, 1, \ldots, N-1$, in Eq. 3.10 is defined as:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn}, \quad k = 0, 1, \ldots, N-1 \tag{3.10}$$

where $X[k]$ represents the spectral components of the sequence $x[n]$ and j is the imaginary unit. The FT allows a signal to be transformed from the time domain to the frequency domain, highlighting the frequencies present and their amplitudes. In signal processing, it is common to use Power Spectral Density (PSD) to estimate the power spectrum of a signal. However, spectrum estimation using FT in real signals can be affected by spectral *leakage* and noise. To reduce these effects, Welch's method is used, a spectral estimation technique that improves the quality of the mean estimate by reducing the variance.

The Welch method consists of the following main steps:

- **Splitting**: The signal $x[n]$ is split into an overlapped segment of L, and in general the overlap is of 50%.
- **Windowing**: Each segment is multiplied by a window $w[n]$, such as the Hamming or Hanning window, to reduce edge effects.
- **FT computation**: For each window, FT is computed.
- **Average**: The power spectral densities of all segments are averaged to obtain a more stable estimate.

The output of Welch's method [99], described in Eq. 3.11 is an esteem of PSD, expressed as:

$$\mathrm{PSD}(f) = \frac{1}{K} \sum_{i=1}^{K} \frac{1}{L} \left|\mathrm{FT}_i(f)\right|^2 \tag{3.11}$$

where K is the number of segments and L is the length of the segment. This approach reduces the variance of the estimator at the expense of spectral resolution.

An example of the spectrum of a signal is shown in Fig. 3.15, computed on a public dataset [86].

It is also possible to apply these transformations to the signal decomposed via the above-mentioned techniques, and to visualise the representation via time, frequency or time-frequency diagrams, that allow to compare different conditions in experimental study. An example of the

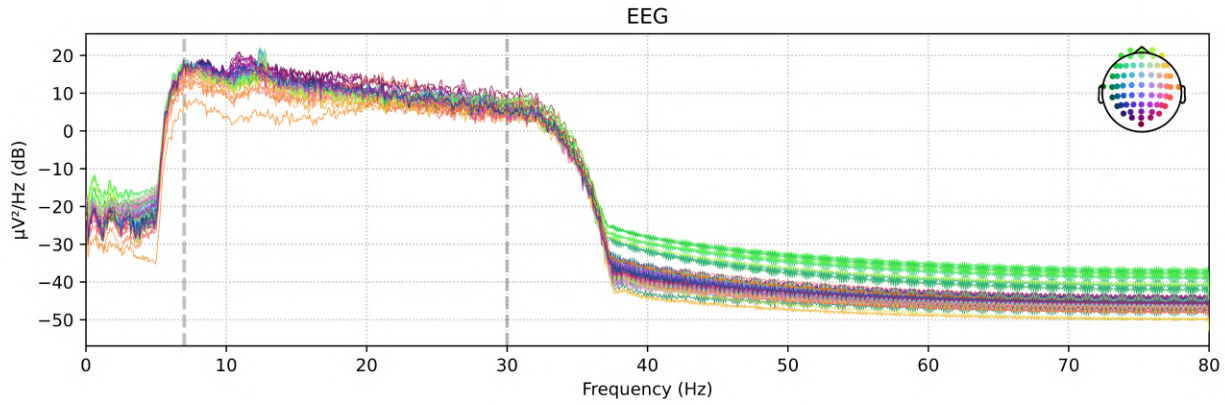


Figure 3.15: The power spectrum of EEG signal after applying a bandpass Finite Impulse Response (FIR) filter between 7 and 30 Hz. On the upper right corner, the electrodeds position coded with the same color of each spectrum of the image. The dotted line represent the bandpass filter.

time-frequency representation is shown in Fig. 3.16. In particular, the time-frequency representation in Fig. 3.16 is Event-Related Spectral Perturbation (ERSP) which is a measure of how the power spectrum of a signal changes over time in response to an event [100]. A widely used representation, such as the one visible in Fig. 3.17 is the topographic representation on the scalp: each channel has an associated value, in this case the PSD for an epoch, and the colours encode the value of the PSD. The study using the ERSP is the basis of Event-Related Desynchronization (ERD)/Event-Related Synchronization (ERS) analysis: in the literature, ERD refer to a decrease in the power of specific frequency bands, which occurs during intense cortical activity and reflects the functional activation of brain regions. On the other hand, ERS is an increase in the power of specific frequency bands, indicating a return to baseline activity or increased synchrony between neurons and often occurs after completion of a task or in resting states. Finally, maps are created by interpolating those values over the entire scalp surface [101] with several different interpolated functions, and the cubic spline is the most common function used. The topographic representation is called Topographical Event-Related Spectral Perturbation (Topo-ERSP). Although classical EEG interpretation considers each region of the brain individually using many features, for example, the PSD of each channel, it is useful, besides single activations, to take into account also the reciprocal relationship between channels [102]. The study of connectivity, also known as connectome, is referred to the analysis of the information flow between brain regions, without information about the directions and this could be analyzed into two main classes: Power-based (EEG) (PoB) or Phase-based (EEG) (PhB). PoB is based on correlatin time-frequency power between electrodes, instead of PhB analyse the distribytuion of phase angles differences over electrodes based on the idea that if two neuronals population are connected, then their oscillatory processes are synchronized in phase [103]. The connectivity could be represented in two ways: Functional Connectivity (FC) and Effective Connectivity Effective Connectivity (EC). FC is referring to studying the information flow between brain regions without information about the directions; EC takes into account also the direction. EC is a more informative analysis, but it is more difficult to perform due to brain anatomy, and it could be very imprecise [103]. EC is usually analyzed by using Granger Causality or Transfert Entropy [102]. Connectivity could be analyzed with mathematical tools divided into two main classes: PoB and PhB. PoB uses the correlation of time-frequency power between electrodes,

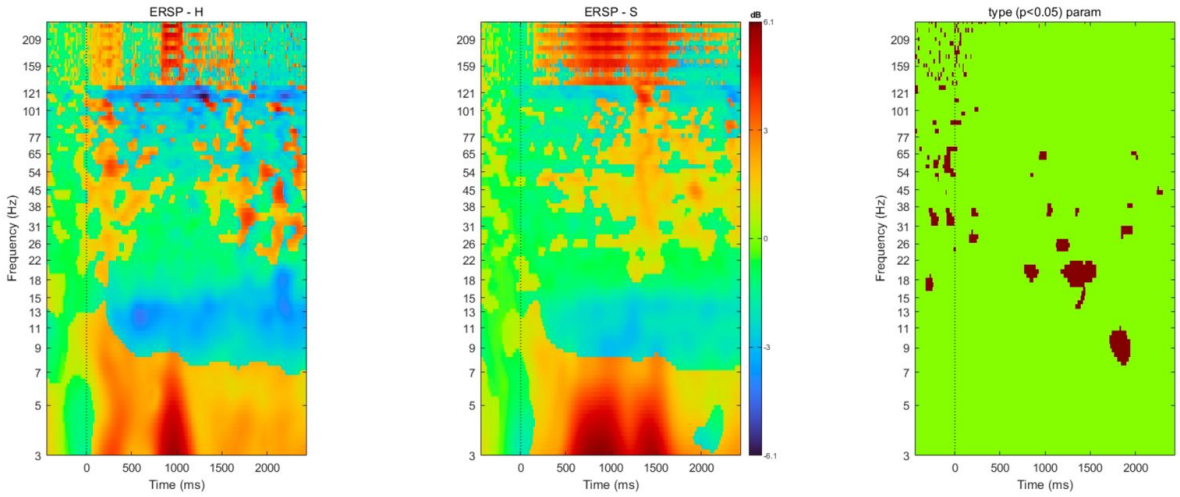


Figure 3.16: Example of time-frequency representation(ERSP) of Fp1 channel in two different experimental condition: on the left the condition H, in the center the condition S, and on the right the result of the p-value for t-test between the two conditions. The data cames from the experiment shown in Chapter 6.

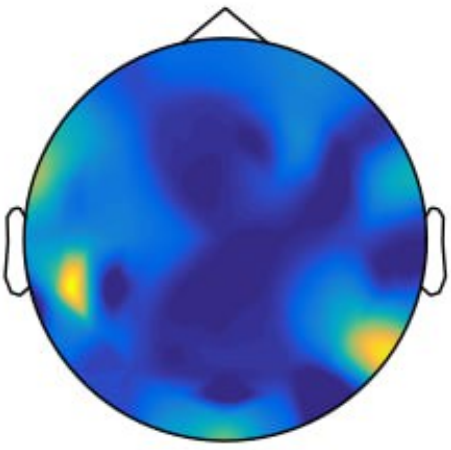


Figure 3.17: Example of topomap representation called Topo-ERSP. The data cames from the experiment shown in Chapter 6.

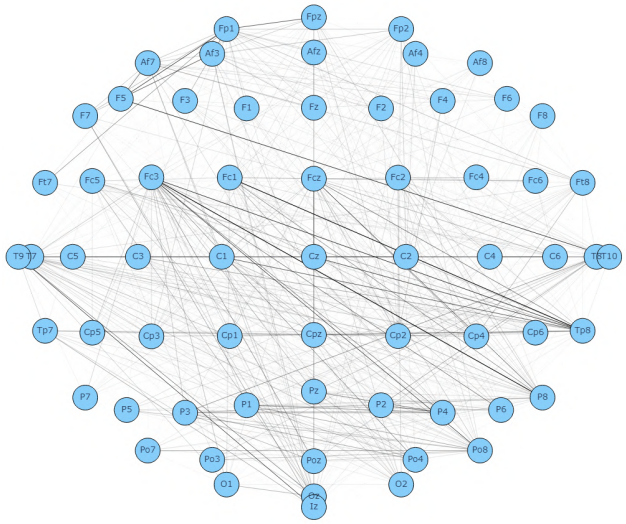


Figure 3.18: A graph-based visualization of the EEG FC based on the PLI. The data cames from the experiment shown in Chapter 6. Image generated with EEGraph library [112].

PhB analyzes the distribution of phase angle differences over electrodes based on the idea that if two neuronal populations are connected, then their oscillatory processes are synchronized in phase [103]. For this reason, phase angle differences are very important for discovering the interactions among different brain regions [104, 105]. The Phase Lag Index (PLI) [106] is a parameter that is able to calculate the asymmetry of the distribution of the phase difference between two signals, which is much less affected than the phase coherence parameter by the influence of common sources and active reference electrodes. Studying connectivity is very informative, unless there is a lack of directionality information and the volume conduction problem: this is a confunding phenomenon that can lead to serious connectivity results for two mainly reasons: the sources in the brain generate electromagnetic fields that are measured by more than one EEG electrode and this introduces spatial autocorrelation at the electrode level [103]. Moreover, the electrical fields spread laterally thought head tissues and this spread to neighboring electrodes. This problem lead to a not easy interpretation of brain localization based on electrode data, and this could create a confounding for many connectivity analysis. This leads to difficulty in interpretation: with volume conduction effects, two electrodes may reflect true connectivity between different brain regions or may reflect the same activity from the same brain source [103, 107]. Furtherore, volume conduction [104] could confound brain connectivity in sensor space and its recognition already difficult open problem due to inverse problem [107]. The study of connectivity depends on signal phase interaction and it can be very informative: 0 or π phase lag is commonly considered a volume conduction consequence, and for this is useful to be insensitive to volume conduction. Moreover, it is well demonstrated that PLI reduces bias due to spurious results by volume conduction with respect to other measures. To analyze the connectivity and the relationship between electrodes, that reflect the connection from two or more brain regions using a multi-channel EEG, one of the most common indexes is the PLI as measure for the connectivity using imaginary part of the signal after the Hilbert Transform. The PLI is useful because it is insensitive to volume conduction because it considers the source space rather than the sensor space [108–111], and has a high reliability with respect to the phase lock value and the synchronization similarity. An example of the PLI is shown in the Fig. 3.18. It is also possible to study the source by means of signal source reconstruction algorithms, using the

subject's anatomy or by using standard brain models that reproduce the tissues and anatomy of the subject. The two best-known methods are the Boundary Element Method (BEM) and Finite Element Method (FEM) and allow the electrical potential measured on the surface of the scalp to be calculated from neural sources within the brain. These consider the signal-generating sources as magnetic dipoles and attempt to identify their origin on the basis of anatomical knowledge, estimating their position with a certain error. For example, Low Resolution Electromagnetic Tomography (LORETA), standardized Low Resolution Electromagnetic Tomography (sLORETA), and Exact Low Resolution Electromagnetic Tomography (eLORETA) are prominent algorithms used in the field of source localization of EEG. These algorithms aim to reconstruct the spatial distribution of electrical activity in the brain from scalp EEG recordings, which are inherently limited by the volume conduction effect, where electrical signals spread out as they travel through the head [113].

3.3.3 Deep Learning

Artificial Intelligence (AI) is a complex and evolving technology, difficult to define precisely and to accept universally. There are various definitions of AI, from the broader ones that equate it to the use of algorithms to the more restrictive ones that consider it as the complete imitation of human intelligence by computers. However, most definitions agree that AI implies the ability of machines to imitate or simulate complex human skills, such as learning, problem solving and decision making [114]. Among the many branches of AI, Machine Learning (ML) is the generic set of techniques that build a predictive model from data. Among the methods of supervised ML most common include the K-nearest Neighbors (KNN), Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), and for unsupervised one of the most common method is K-means (K-means). Supervised methods are methods that use the labels of the data provided for training during the training phase; unsupervised methods, on the other hand, do not know the labels of the training data and clusters the data on the basis of the features of the data. SVM is a supervised ML algorithm that can be used for both classifying and regression. It aims to find the best hyperplane that divides the data into different classes, finding the maximum possible distance of the margin between the hyperplane and nearest data points of each class. LDA is a method for dimensionality reduction that detect the linear combinations of features to maximize the separation between classes using the projection of the data into a space where the distinction between classes is at maximum. K-means is an unsupervised clustering algorithm. It organises the data into K clusters with K to be chosen before training. It iteratively assigns the data to the cluster with the closest centroid and updates the centroids, proceeding to convergence or a maximum number of iterations. KNN is a non-parametric ML technique that assigns a data point to the predominant class among its K-neighbours in the feature domain. The distance between points is evaluated using metrics such as Euclidean the value of parameter k greatly influences the results. Deep Learning (DL) constitutes a segment of ML that draws inspiration from the human brain's architecture and function, utilizing complex algorithms and multi-layered neural networks to process and learn from data.

The *Universal Approximation Theorem* [115–117], is a series of mathematical proofs that a feedforward neural network with at least a single hidden layer can approximate a large number of functions with an arbitrary degree of accuracy. The conditions that must be fulfilled are: (a) possess a sufficient number of neurons, (b) that the activation functions are chosen appropriately (i.e. nonlinear) (c) the function to be approximated must be continuous defined over a compact

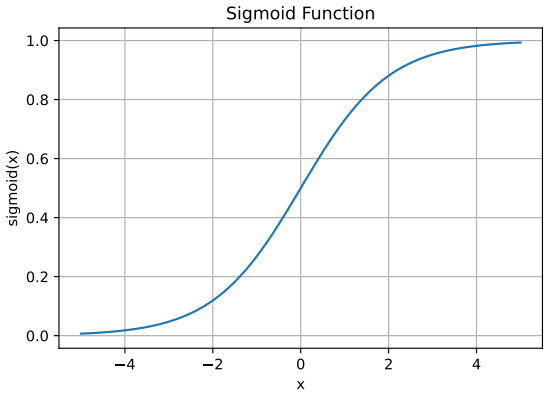


Figure 3.19: The sigmoid function.

subset of $\mathbb{R}^n$. In particular, [115] proves that any continuous function defined over a closed and finite interval can be approximated arbitrarily well by a linear combination of simpler functions, and [116] proves rigorously that multilayer feedforward networks are a class of universal approximators.

In this system, inspired by biological model, a neuron is a computing unit processing weighted inputs to produce an output. Formally, the output y of a neuron can be expressed as the Eq. 3.12

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{3.12}$$

where x_i represents the inputs, w_i the associated weights, b the bias, and f is a nonlinear activation function such as the sigmoid, shown in Eq. 3.13 and shown in Fig. 3.19.

$$f(z) = \frac{1}{1 + e^{-z}} : \tag{3.13}$$

The Perceptron [118] is a model that can have two different activation functions: a threshold or continuous. There exist different layers typically and the neurons of a specific layer are connected with the next layer. A weight is associated with each connection. Backpropagation is an algorithm used to optimize the weights w_i and the bias b in neural network models. It relies on minimizing a cost function $J(w, b)$, such as the mean squared error, using gradient descent. The algorithm consists of two stages: forward and backward propagation. In forward propagation, the model calculates the output y_{pred} from the input, with the error assessed by contrasting it with the true value y_{true}. In the backpropagation step, the partial derivatives $\frac{\partial J}{\partial w_i}$ and $\frac{\partial J}{\partial b}$ are computed using the chain rule. In the case of a single weight w_i, the gradient is given by the Eq 3.14:

$$\frac{\partial J}{\partial w_i} = \frac{\partial J}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial w_i} \tag{3.14}$$

where z is the net input to the neuron. Then, the weights are updated according to the rule shown in Eq. 3.15:

$$w_i \leftarrow w_i - \eta \frac{\partial J}{\partial w_i} \tag{3.15}$$

where η is the learning rate. This iterative process allows the network to converge toward an optimal weight configuration based on the training, validation and test sets. In Fig. 3.20, a classic

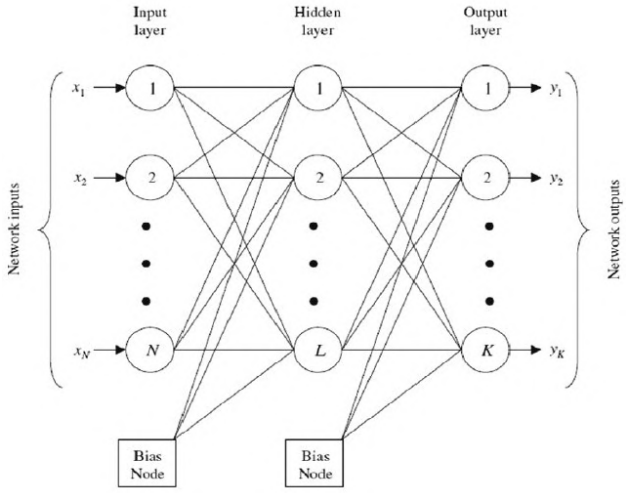


Figure 3.20: A diagram of a classic ANN. Image from [119].

Artificial Neural Network (ANN) architecture is shown. The fully connected ANN is the basic AI algorithm and of the DL: each neuron at one level is connected to all neurons at the subsequent level. This dense interconnection allows for the representation of highly complex data in a mathematical manner. As shown in Eq. 3.16, for a fully connected layer, the outcome of one neuron is calculated by:

$$y = f(w_1x_1 + w_2x_2 + \cdots + w_nx_n + b) \tag{3.16}$$

where:

- y is the output of the neuron,
- f is the activation function (e.g., ReLU, sigmoid, tanh),
- $x_1, x_2, \ldots, x_n$ are the inputs to the neuron (i.e., the outputs of the neurons from the previous layer),
- $w_1, w_2, \ldots, w_n$ are the weights of the neuron,
- b is the bias of the neuron.

A fully connected layer first executes a linear transformation on the inputs, succeeded by a nonlinear activation function. This linear-nonlinear transformation synergy allows neural networks to estimate intricate functions. An explanatory diagram of the functioning of ANN is shown in Fig. 3.21. Among the architectures of DL, Convolutional Neural Network (CNN), Long Short Term Memory (LSTM),a type of Recurrent Neural Network (RNN), and Convolutional LSTM Neural Network (ConvLSTM) are DL architectures used for processing data from EEG. This architecture is efficient in case of matrix-organized data (e.g. image) analysis and pattern recognition by the extraction local features due to convolution filters. This methods, are appropriate for image classification and semantic segmentation [120]. A convolutional layer in a CNN performs a this mathematical operation, a convolution, between a *filter* (or *kernel*) and an input (*feature map*), followed by the addition of bias and the application of an activation function. Mathematically, the output of the

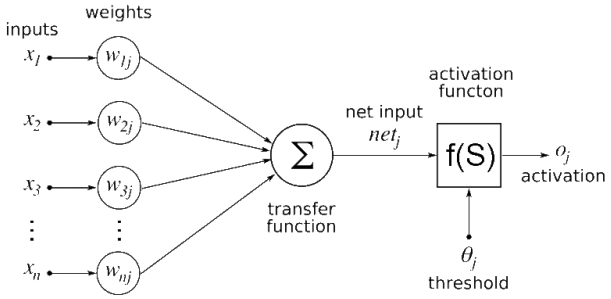


Figure 3.21: ANN diagram. (Author: "Chrislb", CC-SA).

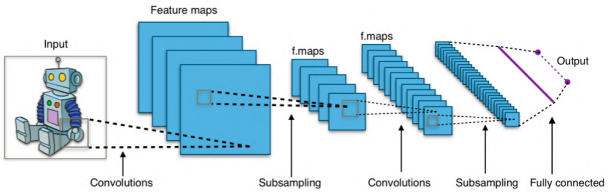


Figure 3.22: CNN diagram. (Author: "Aphex34", CC-SA).

layer is described as follows in Eq. 3.17:

$$Y[i,j] = \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} X[i-k_1, j-k_2] W[k_1, k_2] \tag{3.17}$$

Where $Y[i,j]$ represents the output pixel at position (i,j), X is the input image, and W is the convolution filter or kernel with dimensions $K_1 \times K_2$. This operation is performed by sliding the filter across the input image, calculating the weighted sum of the overlapping pixels. An illustrative schematic of how CNNs works is given in Fig. 3.22.

The CNN have been used with success in many domains, including Computer Vision (CV) and speech recognition, where their architecture allows them to manage complex and high-dimensional data, and are inspired by the biology of visual cortex [121]. The CNN are designed to mimic the way the human brain processes visual information, using layers of neurons that apply convolution operations on input data, such as images or time sequences [54]. They are also able to extract relevant features independently through the convolutional layers, compared to traditional image processing methods. RNNs are a type of ANN designed to process data sequences, such as time signals, text, or time series. Their structure allows them to retain information about past context through a feedback mechanism in the hidden nodes. Unlike traditional neural networks, an RNN uses the same weight W for all time steps, enabling the modeling of sequential dependencies. The update formula for the hidden state h_t in an RNN [122] is show in Eq. 3.18:

$$h_t = f(W_h h_{t-1} + W_x x_t + b) \tag{3.18}$$

where:

- x_t is the input at time t,
- h_{t-1} is the previous hidden state,
- W_h and W_x are the weights,

- b is the bias,
- f is an activation function, typically tanh or ReLU.

and the output y_t at each instant is computed by using the Eq. 3.19:

$$y_t = g(W_y h_t + c) \tag{3.19}$$

where W_y is the output weight matrix, c is the output bias, and g is activation function. RNNs face the vanishing gradient problem, which limits their ability to learn long-term dependencies. To address this issue, LSTMs were developed. LSTMs [122] are a type of RNN designed to capture long-term dependencies. Their architecture introduces memory cells and three types of gates: input, forget, and output. This mechanism allows the control of the information flow through the network. Due to their ability to retain information over time, the LSTM they are ideal for capturing temporal dependencies between signals. The updating of a LSTM can be formally described by the following Eqs. 3.20, 3.21, 3.22, 3.23, 3.24, 3.25:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{3.20}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3.21}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3.22}$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{3.23}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{3.24}$$

$$h_t = o_t \odot \tanh(C_t) \tag{3.25}$$

Where: (f_t): forget gate. This "gate" decides which information from the cell memory (C_{t-1}) should be forgotten. It takes as input the previous hidden state (h_{t-1}) and the current input (x_t), processes them through a linear transformation (multiplying by the matrix (W_f) and adding the vector (b_f), and applies the sigmoid function (σ) to produce a value between 0 and 1. A value close to 0 indicates that the information should be forgotten, while a value close to 1 means it should be retained. (i_t): input gate. This "gate" decides which new information should be stored in the cell memory. It works similarly to the forget gate, taking (h_{t-1}) and (x_t) as inputs, applying a linear transformation (W_i) and (b_i), and then applying the sigmoid function. $(\tilde{C}_t)$: candidate cell state. This represents the new information that is a candidate for being stored in the cell. It is calculated by applying a linear transformation (W_C and b_C) to (h_{t-1}) and (x_t), followed by the hyperbolic tangent function (tanh). (C_t): cell state. This is the internal state of the LSTM cell, a form of "long-term memory". It is updated by combining the previous information (C_{t-1}) filtered by the forget gate (f_t) with the new candidate information $(\tilde{C}_t)$ filtered by the input gate (i_t). (o_t): output gate. This "gate" decides which part of the cell state (C_t) should be exposed as output. It works like the other gates, taking (h_{t-1}) and (x_t) as inputs, applying a linear transformation (W_o) and (b_o), and then applying the sigmoid function. (σ): sigmoid function. A mathematical function that transforms any value into a value between 0 and 1. It is useful for representing probabilities or the "intensity" of a signal. This function is shown in Fig. 3.19. Due to their structure, LSTM are particularly effective for applications such as speech recognition, natural language processing and

time series prediction. Finally, another architecture relevant to the study of data is the ConvLSTM,

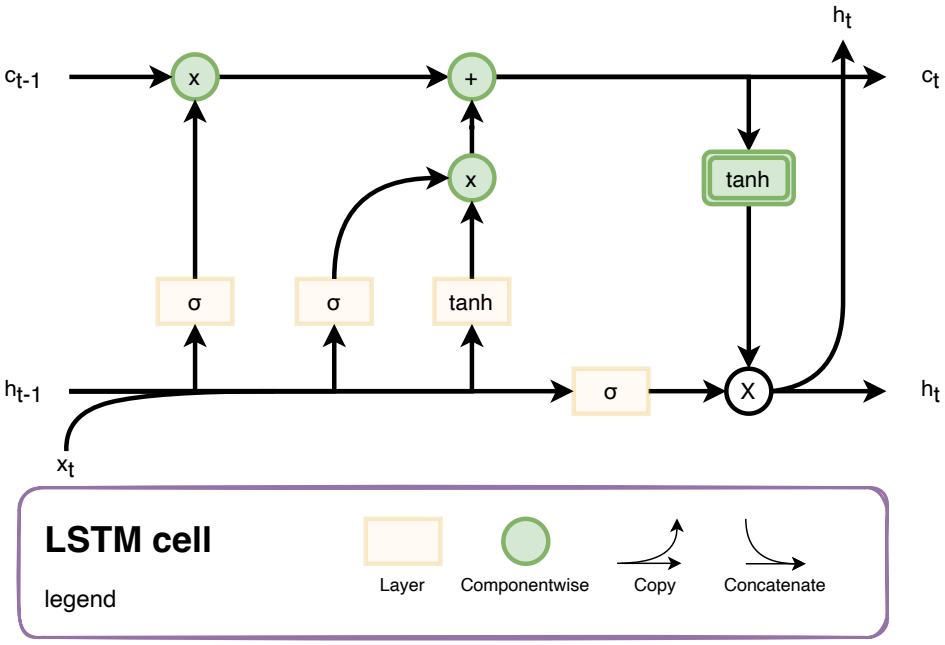


Figure 3.23: Diagram of an LSTM cell.

which combines the spatial capability of the CNN with the temporal capability of the LSTM, thus allowing for complex management of both spatial and temporal characteristics in the EEG data. The ConvLSTM was introduced to address the limitations of traditional LSTM, which are unable to capture spatial information effectively. In a prediction context, for example, the ConvLSTM is capable of analysing sequences of images or spatial data, while maintaining the temporal memory required to make accurate predictions. A diagram of the ConvLSTM layer is shown in Fig. 3.24. For evaluating the model, Accuracy, Precision, Recall, and F1 metrics are used. They are based on True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TPis the number of samples correctly classified as positive, TNis the number of samples correctly classified as negative, FPis the number of samples wrongly assigned to the target class and FN is the number of samples wrongly classified as negative over all positives. Based on these key values, the confusion

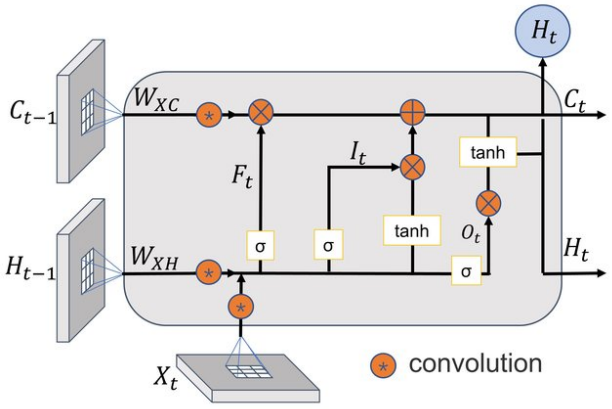


Figure 3.24: The Conv-LSTM layer structure. Image from [123].

matrix is created. Assesses a model's performance by showing the number of correct and incorrect predictions for each class. For binary classification problems, the confusion matrix is a 2×2 table. Each cell represents the number of instances where the actual and predicted classes intersect. The confusion matrix can be represented as table shown in Tab. 3.1.

Actual/Predicted	Positive (Predicted)	Negative (Predicted)
Positive (Actual)	TP	FN
Negative (Actual)	FP	TN

Table 3.1: Binary confusion matrix with TP, TN, FP), and FN

For multiclass problems, the confusion matrix is expanded to an $n \times n$ table (where n is the number of classes). Each cell (i, j) represents the number of instances where the actual class is i, and the predicted class is j. The table for multiclass problem, is shown in Tab. 3.2. In the

Actual/Predicted	Class 1	Class 2	$\cdots$	Class n
Class 1	TP_1	$FP_{1,2}$	$\cdots$	$FP_{1,n}$
Class 2	$FN_{2,1}$	TP_2	$\cdots$	$FP_{2,n}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
Class n	$FN_{n,1}$	$FN_{n,2}$	$\cdots$	TP_n

Table 3.2: Multi-class confusion matrix with TP, FP), and FN for each class.

multiclass case, the key values are the following:

- TP_n (TP) is the value in the diagonal cell (n, n).
- FP_n (FP) is the sum of all values in column n, excluding TP_n.
- FN_n (FN) is the sum of all values in row n, excluding TP_n.
- TN_n (TN) is the sum of all values not in row n or column n.

The definitions of the metric are the following:
Accuracy is defined as the ratio of the correct prediction in all classes over the total set. The formula of the accuracy is shown in Eq. 3.26:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.26}$$

Precision is defined as the ratio of the TPpredictions to the total number of positive, $TP + FP$. The formula of the precision is shown in Eq. 3.27:

$$Precision = \frac{TP}{TP + FP} \tag{3.27}$$

Recall is defined as the ratio of the correct prediction in single classes over the total set. The formula of the precision is shown in Eq. 3.28:

$$Recall = \frac{TP}{TP + FN} \tag{3.28}$$

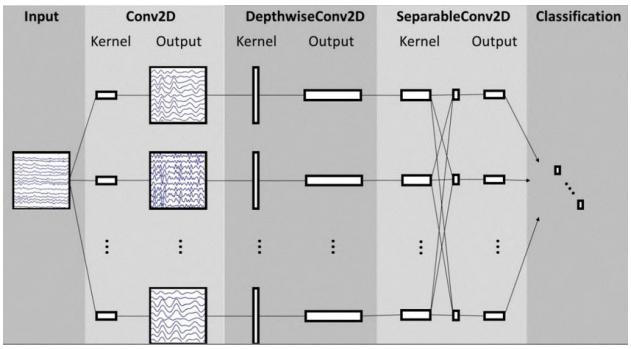


Figure 3.25: EEGnet Architecture. Image from [124].

F1 score represents the power of discriminability between each class taking into account the Precision and Recall metrics of all classes. The formula of the F1 score is shown in Eq. 3.29:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{3.29}$$

3.3.4 Specific DL architecture for EEG data

Some DL architectures have been designed for EEG signal analysis, taking into account the signal characteristics, the spatial distribution of the channels, and the nonlinearities of the signal. Models such as EEGNet (EEGNet), EEG Inception Temporal Network (EEGITNet) and Deep 4-Net (Deep4Net) are examples of DL architectures specifically designed for the classification of these types of signal. Here below, these three architectures used in some works of this thesis are presented.

3.3.4.1 EEGnet

EEGNet [124] is a specialized DL architecture designed for the classification of EEG signals, particularly in the context of BCI. It is characterized by its compact structure CNN, which allows it to effectively process data EEG with a relatively small number of parameters. It employs several convolutional layers, which are competent in extracting features related to time and space of the EEG signals, key factors for correctly interpreting a wide range of cognitive activities with Motor Imagery (MI) [125]. One of the main novelties in EEGNet is the capability to generalize to subjects and tasks, since one of the biggest challenges in EEG-based BCI applications is that the signals of different brains vary considerably [126]. It generalizes through its shallow architecture design to prevent overfitting without compromising on essential feature learning from the EEG data. The architecture in Fig. 3.25.

3.3.4.2 EEG-ITnet

The EEGITNet [127] stands for Electroencephalogram-Inception Temporal Network, which is an advanced approach for MI classification based on EEG signals. This model is designed to address the challenges associated with MI classification by leveraging the power of deep neural networks, particularly CNN and Inception architectures. EEG-ITNet has the special capability of extracting both temporal and spatial features from the EEG signals, enhancing accuracy in MI classification, a critical task for BCI [127]. This architecture consists of a module called *Inception* that is designed

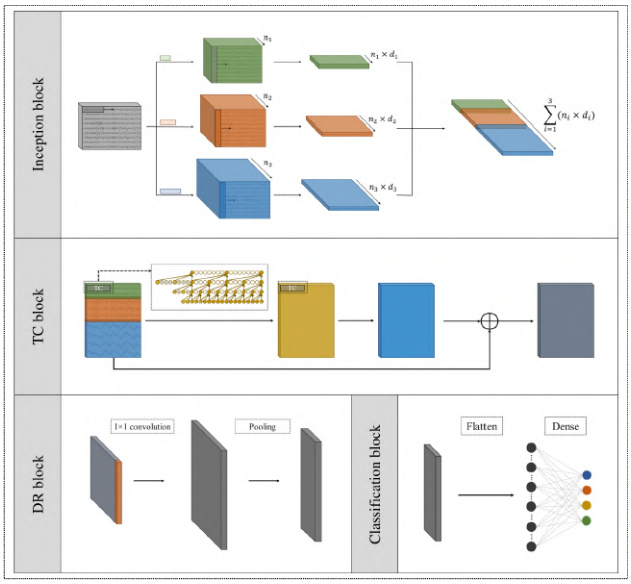


Figure 3.26: EEG-ITnet Architecture. Image from [127].

to enhance the performance of DL models in several applications. Its architecture allows the parallelization of several convolutional filters of different sizes, which enables the network to capture several scales regarding features simultaneously. The Inception module contributes to the network's ability to learn a rich feature representation by integrating convolutional layers with different kernel sizes. This design mitigates the risk of losing important spatial information, which can be an issue in designs that incorporate a standard pooling layer. Another novelty in EEG-ITNet is its explainability, an important factor in the acceptance of DL models in clinical and research usage. It is especially important in the context of BCIs since decisions inferred from EEG signals can have great consequences in terms of neurological disorders treatment and to improve the quality of life of the patient [127]. Fig. 3.26 shows the architecture.

3.3.4.3 Deep4Net

Deep4Net [128] is a CNN architecture with multiple convolutional and pooling layers, capable of extracting even the most intricate features in raw EEG data. Depth and complexity in architecture provide the main salient features to Deep4Net, enabling it to capture a wide range of temporal and spatial characteristics from the EEG recording. Furthermore, the real-time processing ability of the architecture with respect to EEG data is especially suitable for applications that need immediate feedback, such as neurofeedback therapy and cognitive training. Without the softmax layer and by utilizing loss functions of Mean Squared Error (MSE), researchers have prepared Deep4Net for regression tasks, too, extending their work from mere classification problems [129]. The architecture is displayed in Fig. 3.27.

3.4 EEG-based Brain-Computer Interfaces

The term BCI refers to a set of techniques, tools, algorithms, and paradigms that allow the signals produced by CNS to be used to communicate with the external environment through the use of electronics and computer science [130–132]. It is described as a communication and control

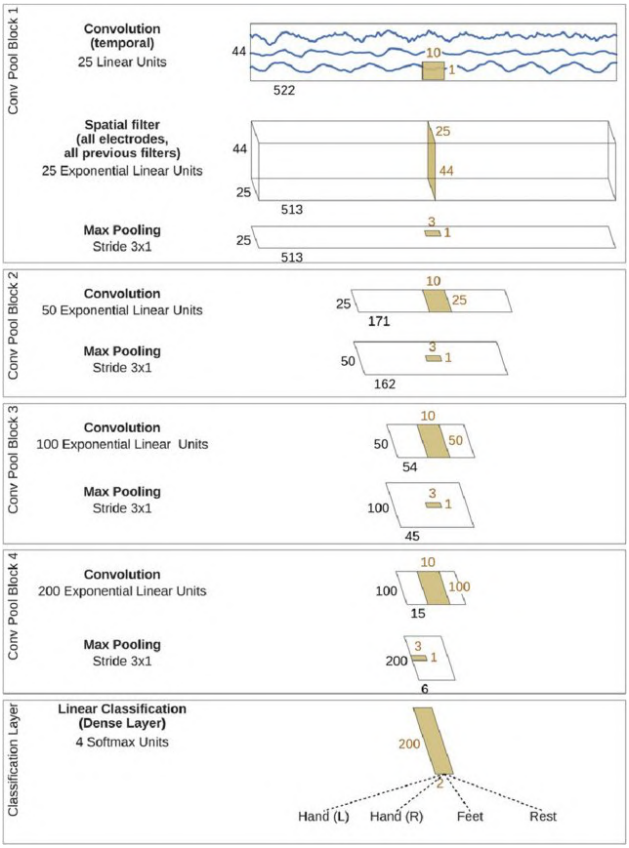


Figure 3.27: Deep4Net Architecture. Image from [128].

system that allows for immediate interaction between the human brain and external devices. The system is designed to monitor activity within CNS and transform this information into an artificial output. This output serves to substitute, repair, increase, support, or improve the natural output produced by CNS [131]. The user's intent of a BCI, manifested through brain signals, is converted by the BCI system into a desired outcome: computer-mediated communication or control of an external [133] device. The technology of BCI is based on studying, cleaning, transforming, analyzing, and classifying the brain, metabolic or electrical signal, and converting it into an output that enables robotic movement, the initiation of a process or speech using a sound device [132]. Among the many uses that can be made of BCI are to assist people with severe disabilities, motor or cognitive, safety at work, and to improve performance in a work or play context. In the most direct application, BCI allows a person to communicate with or control the outside world without using conventional neuromuscular pathways. This allows people with severe physical disabilities both from the point of view of the effector (arm, hand, etc.) and from the point of view of nerve signal transmission from PNS, to use the nerve impulse of CNS again to perform an action. This is even more evident with people in "locked-in" syndrome: this is a rare neurological condition characterized by complete paralysis of voluntary muscles, with the exception of vertical eye movements and blinking. Patients with this syndrome are fully conscious and capable of seeing their surroundings, but cannot communicate or move due to damage to motor pathways in the brain stem.

A key aspect of BCI is their ability to improve the quality of life of users by offering new ways of interacting with their environment. The BCI can be divided into two categories [134]:

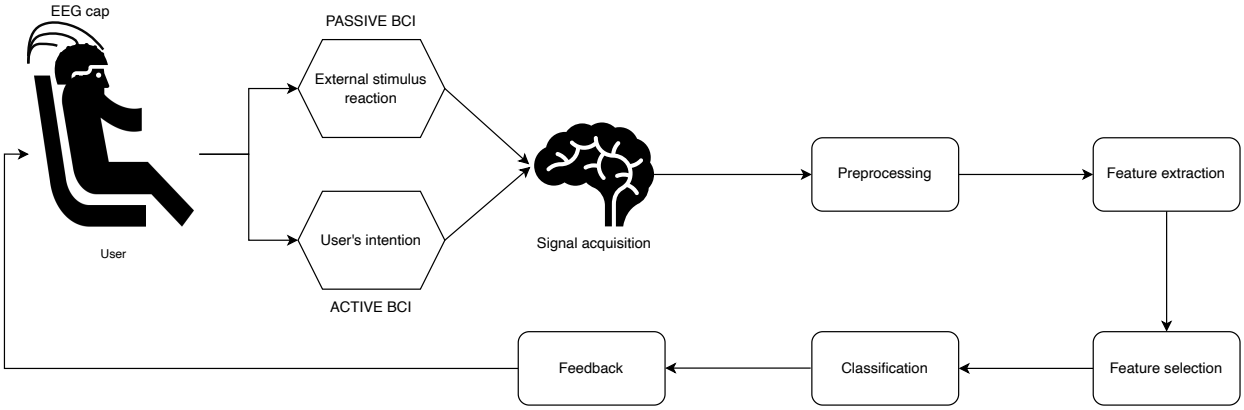


Figure 3.28: An ideal BCI that contains both *passive* and *active* mental-state detection.

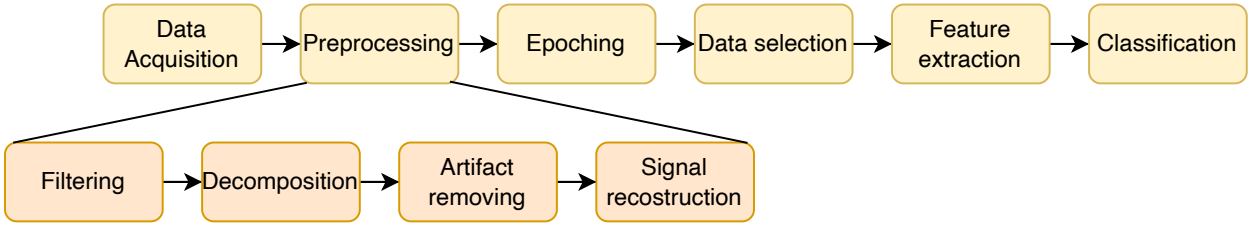


Figure 3.29: A standard pipeline for EEG data analysis.

Passive: is a recognition system for mental states induced by external stimulation (images, sounds, etc.), such as the case of [97] or [135]. An example in this case is the use of ERP or, more generally, of EP. In passive paradigms, individuals do not actively control brain activity; rather, it is used primarily to monitor their psychophysical condition [134].

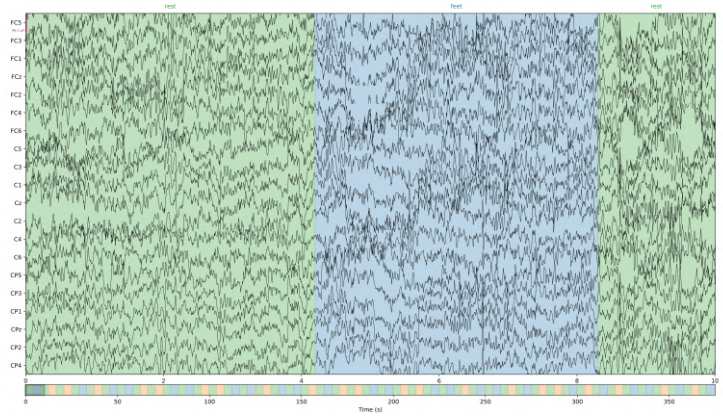
Active: is a recognition system of internal mental states, not induced by any kind of exogenous situation, such as MI [136], or a self-induced emotion [137]. In the context of active paradigms, the individual intentionally modulates brain activity to handle an external interface, independent of external stimuli [134].

It is therefore possible to define that *an ideal BCI is composed of both an active and passive system of detection of mental states*, as shown in Fig. 3.28.

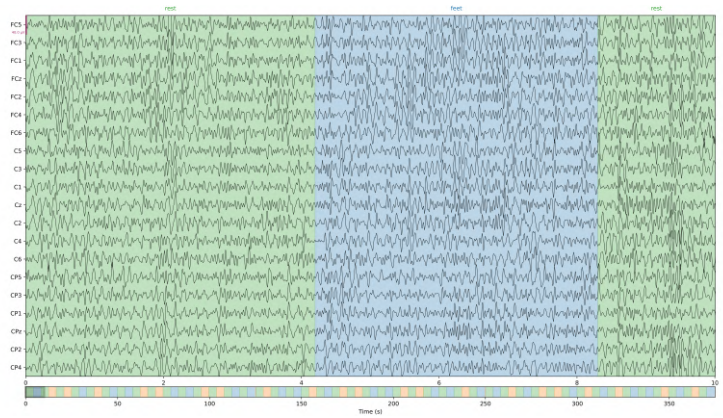
3.4.1 Pipeline for BCI

The data analysis pipeline EEG, illustrated in Fig. 3.28 and more in detail for data preprocessing in 3.29, includes basic steps that can be summarized as follows:

1. **Signal acquisition:** The signal is acquired via a device EEG as a result of a stimulus or cognitive action performed;
2. **Preprocessing:** preprocessing includes cleaning the signal of physiological or instrumental and environmental artifacts using filtering, decomposition and automatic systems;
3. **Feature extraction:** feature extraction can be done on the basis of previous studies or automatically by means of AI algorithms, and this step can be task-dependent;
4. **Feature selection:** relevant features are selected for classification;



(a)



(b)

Figure 3.30: Example of (a) raw EEG signal and (b) filtered signal between 1 a 30 Hz. The different background colours in (a) and (b) represents different mental tasks (green: rest; blue: feet movement)

5. **Classification:** Using a classification algorithm, the signal epoch class EEG is determined;
6. **Feedback:** the system performs the action associated with the mental task or sends feedback regarding the detection of the mental state.

First, signals are acquired through electrodes placed on the scalp, which record changes in electrical potential generated by neuronal activity. This step is crucial, as the quality of the collected data directly influences the performance of the classification models [138]. Once acquired, signals need a thorough preprocessing step, which includes the removal of artifacts, as discussed in this chapter. A notch filter is generally applied to remove interference due to current, and a band-pass filter to exclude the frequencies that are not in the scope of the research question [53]. An example of a bandpass filter can be seen in Fig. 3.31, and an example of an ideal filter is defined in Eq. 3.30. In Fig. 3.30, an example of the filter applied to EEG data from a public dataset [86] is shown.

$$H(f) = \begin{cases} 1, & f_1 \leq f \leq f_2, \\ 0, & f < f_1 \text{ o } f > f_2. \end{cases} \tag{3.30}$$

There are several standardised preprocessing pipelines [140, 141] such as early stage PREProcessing (PREP) [142] that allows for data cleaning EEG and control of data cleaning steps. In particular, early stage PREProcessing (PREP) can be used to effectively eliminate noise from the d.c., to

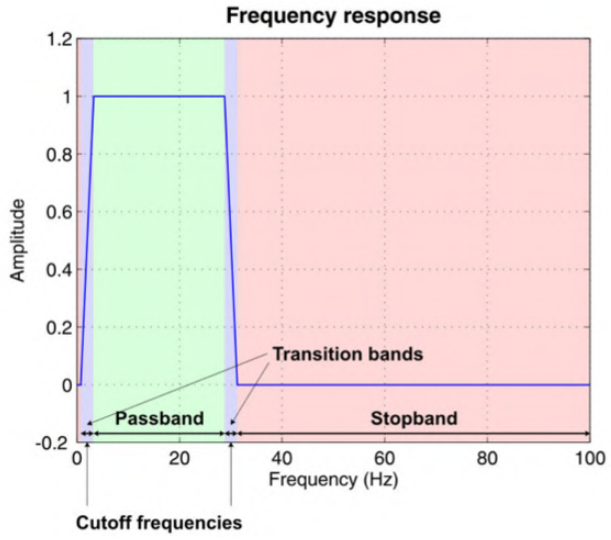
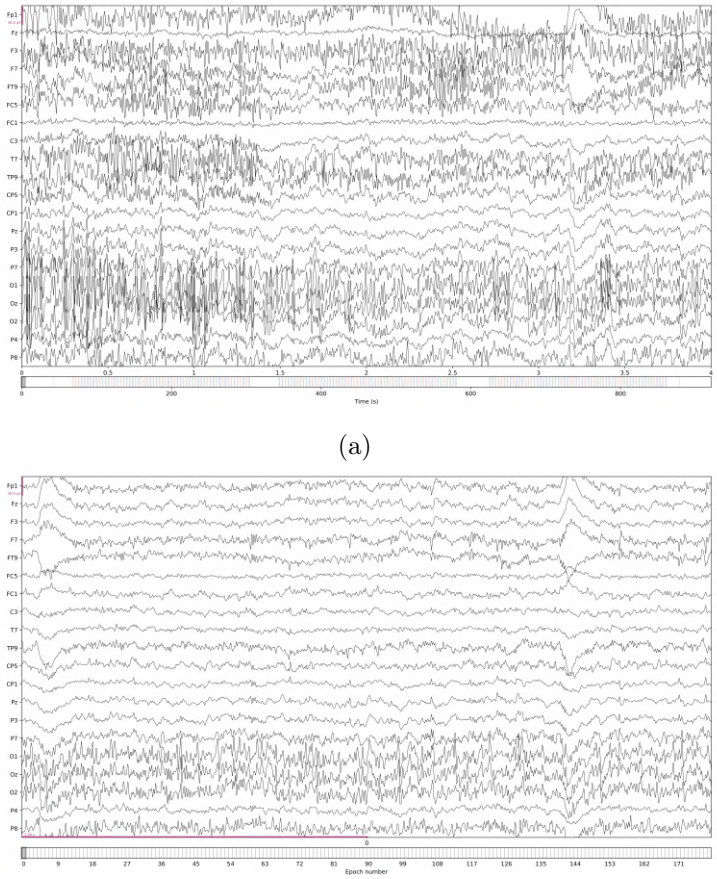


Figure 3.31: A band-pass filter. Image from [139].

perform robust re-referencing, and to detect, remove and interpolate very noisy channels. An example of the PREP results is shown in Fig. 3.32 using the data illustrated in Section 6.4.

The study of Delorme [140], examines preprocessing techniques, textitasizing the significance of high-pass filtering as a critical initial step, while also highlighting the cautious application of referencing to preserve data integrity. Subsequently, feature extraction is crucial to reduce the dimensionality of the data and to identify significant patterns. Feature selection is a critical step that aims to identify the most relevant variables for the classification task, thus reducing the risk of overfitting and improving model accuracy [143]. Various feature extraction techniques, such as CSP, PCA and FT, are frequently employed to obtain useful representations of EEG signals [144], together with other methods of signal decomposition and transformation outlined above. Finally, the classification phase uses various machine learning and DL algorithms to categorise the signals according to the patterns identified, evaluated with several metrics accoring to the threshold of the statistical significance. In [145], the authors present results from simulations to determine what threshold should be taken into account in experiments with EEG data to consider a result above the chance threshold. The classification with ML algorithms of signals is at the heart of both offline and online passive and active BCI. However, in recent years, DL models have gained popularity due to their ability to learn complex and nonlinear representations from raw data. The best-known architectures in the scientific literature such as EEGNet, EEGITNet and Deep4Net, described above, based on convolutional and/or recurrent layers, were built specifically to process EEG data and have been described extensively in the section 3.3.3.



(a)

(b)

Figure 3.32: (a) raw EEG signal and (b) the EEG signal after PREP cleaning process.

Part II

Specific Studies and Related Computational Models

Chapter 4

Motor Execution

The results of this chapter are published in:

E. Mattei, D. Lozzi, A. Di Matteo, C. Manes, F. Mignosi, M. Polsinelli, G. Placidi, "**Analysis of the EEG resting-State Signals for BCI**", *9th Graz Brain-Computer Interface Conference 2024*, pp. 355–359, 2024. DOI: 10.3217/978-3-99161-014-4-062

E. Mattei, D. Lozzi, A. Di Matteo, M. Polsinelli, C. Manes, F. Mignosi, G. Placidi, "**Deep Learning Architecture analysis for EEG-Based BCI Classification under Motor Execution**", *2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 549–555, 2024. DOI: 10.1109/CBMS61543.2024.00097

E. Mattei, D. Lozzi, A. Di Matteo, A. Cipriani, C. Manes, G. Placidi, "**MOVING: A Multi-Modal Dataset of EEG Signals and Virtual Glove Hand Tracking**", *Sensors*, vol. 24, no. 16, pp. 5207, 2024. DOI: 10.3390/s24165207

The presented dataset is published in:

Mattei, E., Lozzi, D., Di Matteo, A., Placidi, G., Cipriani, A., & Manes, C. (2024). "**MOVING: a Multi-MOdal dataset of EEG signals and VIrtual Glove hand trackING [*Dataset*]**". Zenodo. https://doi.org/10.5281/zenodo.12804784

This chapter presents studies on the study of ME as a first step for the encoding of neural signals.ME was selected due to its well-documented neurophysiological information and its ability to ensure an almost definitive ground truth, with movements being controllable visually and by sensors such as cameras or EMG. In particular, in this chapter the studies related to the analysis of the data from EEG during ME will be presented, particularly comparing it to a *resting state*, which is fundamental for the development of active and passive BCI systems, as shown in Chapter 3. The first study proposes a novel architecture composed of multiple concatenated EEGnet models organized in a tree structure, aiming to classify right-hand movement, left-hand movement, and the *resting state*. As a natural continuation of this section, the results are presented on the analysis of three different neural networks and two distinct pre-processing pipelines to determine the best combination to classify EEG movement signals while considering the *resting state*. Finally, an innovative data set has been made available to the scientific community for the analysis of EEG movement data recorded simultaneously with Virtual Glove (VG), a device for acquiring kinematic hand data, developed at the A^2VI-lab, to improve the accuracy of movement or *resting-state* classification by incorporating kinematic data during the training phase. The application of engineering and computer science

to movement analysis can improve the understanding and quantification of gestures by standardizing the acquisition and analysis process. This results in more accurate and objective data, allowing movement trends to be charted and compared with ideal models or other subjects. The evolution of data collection techniques for hand movement, both in motion analysis and in Human-Computer Interaction (HCI), has followed a remarkable trajectory of technological development, and various methods for tracking hand gestures have been developed [146]. In the early stages, hand movement monitoring devices used RGB cameras or depth sensors to record movements in 3D [147]. These systems required advanced algorithms to process spatial data and determine the positions and movements of hand joints over time. However, they had significant limitations, such as the need to accurately calibrate the sensors and difficulties in handling occlusions. The hand and finger movement tracking sensors developed by the company Ultraleap [1] are of a small size (approximately 80 mm x 11 mm x 30 mm). The Leap Motion Controller (LMC) device uses a Infrared (IR) camera with an inverted pyramid-shaped field of view with a volume of 0.22 cubic meters and a height of approximately 600 mm. The pyramid has angles of 150 degrees wide and 120 degrees deep [148–150].

Recent systems have been introduced that combine two orthogonally arranged sensors LMC in a system called VG, minimizing occlusion problems and improving tracking accuracy, making these devices particularly suitable for hand rehabilitation. The development of VG started with the analysis of the two sensors LMC individually and then constructed a monolithic system consisting of two orthogonally arranged LMC for the simultaneous analysis of the hand from two different observation points. An example of VG is shown in Fig. 4.1.



Figure 4.1: A picture of the VG.

Finally, several algorithms and methods were developed that allow the integration of the data from the two sensors to generate a new improved data set resulting from their fusion [151]. The data structure derived from the VG (and consequently also from the individual devices LMC that make up it) includes information regarding the position, velocity, angles of each joint that makes up the virtual hand model.

[1] https://www.ultraleap.com/

In the literature, EEG-based BCI is used to decode a user's intention of movement known as Motor Imagery (MI), and as well as to decode real human movements, Motor Execution (ME), using AI [152]. In general, the study of BCI also focuses on recognizing movements of the hand, upper limb, wrist, and fingers from human neurophysiological signals to control external devices such as robotic arms. To realize a BCI, it is necessary to have a dataset and a model that needs to be trained.

The ME datasets, available in the literature, collect many types of movement, ranging from the most complex to the most intuitive, and DL techniques are applied to classify these data. The EEG Motor Movement/Imagery Dataset [153] was developed by obtaining signals from the BCI2000 64-channel system by choosing 109 volunteers who imagined opening and closing the left or right hand, opening and closing both hands or both feet according to video stimuli, and then they are replicated with the real movement of the subjects. Another dataset is EEG Data for the Voluntary Finger Tapping Movement which is a collection of EEG data acquired during voluntary asynchronous index finger tapping by 14 healthy adults. EEG was recorded using a TruScan Deymed amplifier with 19 channels for three conditions: tap of the right finger, tap of the left finger, and resting state, with a sampling rate of 1024 Hz. Each participant performed 120 tests, 40 for each of the three conditions [154]. In the case of online BCI, in addition to classifying movement, the phase without motion must also be identified, also called the resting state. Some work has been published that refers to the classification of rest for MI. In [155], a hybrid model that combines CNN and transformer architectures, ConTraNet, was implemented using their strengths to improve classification performance in various EEG applications, including categorization of "rest". In [156] the VS-LSTM model was introduced to classify the limb MI using EEG signals, with a particular focus on distinguishing between motor and resting states. In the study by [157], novel approaches for ME/MI classifications were introduced, achieving a high level of accuracy. However, the authors did not take into account the rest period for the classification. In [158], the authors performed the classification on the MI signals, open and closed eyes, using the LDA, Naive Bayes (NB) and SVM classification algorithms achieving an accuracy of 91.18 %, 95.41 %, and 99.51 %, respectively. Furthermore, in [159] the authors classified 2, 3, and 4 classes using EEGnet in MI and rest tasks, reaching and overall accuracy of 82.43 %, 75.07 %, and 65.07 %, respectively. Also in [160], the authors classified the MI signals and rest, reaching an accuracy of 70.64 % in 5 classes. This section shows that most studies conducted on resting state are carried out for MI, leaving out ME. In [161, 162], it can be observed that the two cognitive processes MI and ME are different both anatomically and functionality: in [161], the authors found that during the grasping activity, the PMAs dynamically interplay in different ways, based on the task demands, whereas in [162], the authors found a different activity using fNIRS between ME and MI.

In a recent review [163], many EEG motor-related datasets are summarized and many of them are related to MI tasks, while few are related to ME. The distinction between MI and ME tasks is a well-known concept in scientific literature, and can be identified using neuroimaging methods [162]. Among the ME dataset, the PhysioNet Motor Movement/Imagery Dataset [86] was created by recording signals from the BCI2000 64-channel system from 109 volunteers who imagined moving their hands, feet, or both in response to video stimuli. These signals were then replicated with the actual movements of the subjects. The Upper Limb movements dataset [163] consists of data from 15 healthy subjects, all except one were right-handed. To avoid muscle fatigue, subjects sat in a chair with their right arm fully supported by an exoskeleton. Each subject had two sessions on different days, no more than one week apart. Subjects conducted ME in the first session and MI in the second. Six different movements were made with the right upper limb, beginning in a neutral posture. Additionally, a rest class was recorded in which individuals were advised to avoid moving and keep their starting position.

The EEG BCI dataset [164] was constructed by capturing signals from electrodes F3, F4, FC5, and FC6 from the Emotiv Device, and selecting four individuals to imagine the opening/closing of the right or left hand after the video stimuli for a total of 2400 trials. The EEG Motor Movement/Imagery Dataset [153] was developed by obtaining signals from the BCI2000 64-channel system by choosing 109 volunteer individuals who imagined opening and closing left or right hand, or opening and closing both hands or both

feet according to video stimuli. The BCI Competition III dataset IIIa [165] has four classes (left hand, right hand, foot, tongue) acquired from three subjects by obtaining signals from a 64-channel Neuroscan EEG amplifier for a total of 60 trials per class. The BCI Competition III dataset IVa [166] was recorded with multichannel EEG amplifiers using 64 or 128 channels from ten healthy subjects. Subjects sat in a comfortable chair with arms resting on armrests and imagined left or right hand or foot or imagination of a visual (with eyes open), auditory, or tactile sensation. Moreover, BCI Competition IV dataset 2a [167] and 2b [168] are also available for MI. For the classification, many DL techniques are used. A deep metric model with a triple CNN was used to classify MI of the right or left hand based on incoming visual stimuli; in the absence of a stimulus, it is deemed rest [169]. Electroencephalography Topographical Representation (ETR)-CNN classifies five MI movements and one resting movement using the ETR manipulation on the signals of the BCI Competition IV-2a dataset [170]. Classification Network-based LSTM is applied on the BCI Competition IV-2a dataset and to make the classification more resilient on the dataset's signals, a one-dimensional-aggregate approximation is performed [171], as well as in other BCI domains such as emotions [2, 173]. Regarding ME, the approach used is the same as above for MI. Datasets, available in the literature, collect many types of real movements ranging from the most complex to the most intuitive, and DL techniques are applied to classify them. EEG Data for Voluntary Finger Tapping Movement is a collection of EEG data acquired during voluntary asynchronous index finger tapping by 14 healthy adults. EEG was recorded using a TruScan Deymed amplifier with 19 channels for three conditions: right finger tap, left finger tap, and resting state, with a sampling rate of 1024 Hz. Each participant performed 120 trials, 40 for each of the three conditions [154]. The High-Gamma Dataset is a 128-electrode dataset collected from 14 healthy subjects during four-second trials of executed movements, separated into 13 runs per subject. Movements can be classified into four types: left hand, right hand, both feet, or rest [174].

4.1 Analysis of the EEG Resting-State Signals for BCI

The present work describes the activity of the first stage that focuses on analyzing the response of the rest signal to the real movement of the hand (left and right) for the robotic control arm in rehabilitation [175]. The objective of this research is to determine the optimal method for distinguishing the data EEG from the data EEG from the data of ME by employing a deep learning architecture, EEGnet[176]. This involves decomposing the three-class problem (rest, left, right) into several binary classifications (rest vs left, rest vs right, left vs right) and addressing this through the construction of a *Network Tree* that integrates several EGGnet trained to identify rest and the two movement-related classes, to improve classification performance compared to the EEGnet trained on three classes that showed low accuracy during our previous analysis.

4.1.1 Materials and methods

Data from 105 people from the Physionet Dataset [153] were used. The original data set consisted of 109 persons, but four subjects were excluded because they conducted different numbers of trials with a different sampling rate. Each subject was recorded while performing an ME or MI of opening and closing the right or left hand for 4 seconds. Before each exercise performed or imagined, there was a 4 s rest. Each person was recorded three times, and each recording contained approximately seven movements in each category, plus one rest period for each movement. Of 64 electrodes, only the following 15 were used: Fp1, Fp2, F7, F8, Fz, P3, Pz, P4, O1, O2, F3, F4, C3, C4, Cz. This was done to reduce the number of channels and make the system lighter and more portable, given the possibility of using a cheaper EEG device (such as Enobio 20[2] or other light mobile device EEG. After exclusion, the epochs were extracted using one second of signal, starting from the marker, without any baseline reduction. The signals were downsampled to the data from 160 to 80 Hz and a highpass filter to 1 Hz was used. In Fig. 4.2, the preprocessing pipeline applied to the dataset and the classification stage are shown.

[2]https://www.neuroelectrics.com/solutions/enobio/20

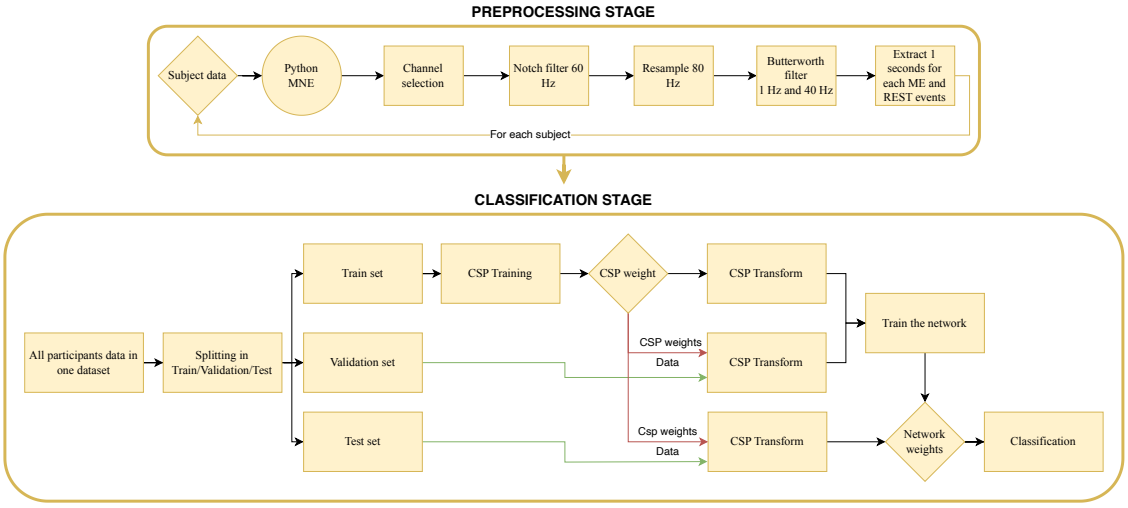


Figure 4.2: The pipeline used for data preparation and classification stages.

Our work aims to create a deep learning model capable of recognizing the movement execution of the right hand, the left hand, and the rest period. For this reason, only the EEG data acquired during the ME conditions were used. Finally, our dataset was composed of 4725 signals under "ME rest condition", 2369 signals under "ME left-hand movement" and 2356 signals under "ME right-hand movement". In each class, the corresponding trials of all selected subjects are grouped. In Fig. 4.3 the split of the train/validation/test sets is shown. The movement decoding process was carried out using the DL architecture EEGnet, and from now on EEGnet [176], trained on the EEG signal from the EEG motor movement/image dataset [153]. The preprocessing and deep learning models were performed using Python with the MNE library [177] for data analysis EEG, PyTorch [178] for deep learning and [179] for the EEGnet architecture. Before training, starting from the original data, the transformation was performed using CSP [98, 180] and used the transformed signal as input for the network. This operation reduced the number of channels from 15 to 2, maximizing the distinguishability between classes. This process was repeated for all the EEGnet used. All networks were trained using a balanced training set, so all classes had the same number of samples in both the training and the test. The training hyperparameters are displayed in the Tab. 4.1, were selected after several experimental trials. It was noticed that signal downsampling did not affect performance, so it was downsampled the signal to improve training speed. The Stochastic Gradient Descent (SGD) was used to improve the convergence of the training, also applying a dynamic Learning Rate (LR), starting from a value of 0.004, which reduced its value automatically when, after 25 epochs of inactivity, it did not improve. The factor of reduction of Learning Rate was set at 0.2. Moreover, a double batch size [181] was used that changes when there was no further improvement of the network also after reducing the Learning Rate. Finally, it was used a decay rate of 0.001 to improve the learning rate of the network.

The signals of rest, execution of open close right-hand (right) and execution of open and close left-hand (left) were considered. As a first step, the EEGnet network was trained in the three classes rest (0), left (L) and right (R), called 0LR, with unsatisfactory results, as shown in the Tab. 4.2 in the last row. Therefore, comparisons were made between the same EEGnet trained in two left-right, left-right, and right-right classes defined as LR, 0L and 0R, respectively. After this analysis, it was observed that the 0LR network did not correctly distinguish the three classes to the LR, 0L, and 0R networks. For this reason, a *Network Tree* based on previously tested EEGnets was developed to improve classification accuracy and analyze the rest signal. This "*Network Tree*", was composed of three different EEGnet [176] individually trained on different subsets of our dataset, to classify different types of signal:

- $EEGnet_{LR}$: the network was able to classify correctly the signals labeled as left or right;
- $EEGnet_{0L}$: the network was able to classify correctly the signals labeled as rest or left;

- $EEGnet_{0R}$: the network was able to classify correctly the signals labeled as rest or right.

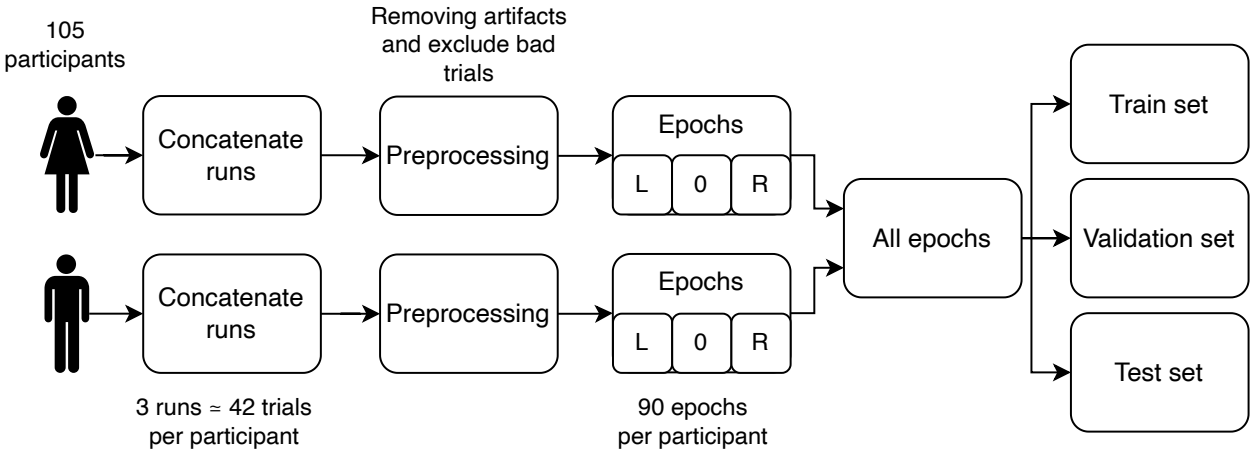


Figure 4.3: The split of the original set in train, validation and test set after preprocessing and epochs extraction.

Hyper-parameter	Value
Sample shape	1, 15, 79
Sampling rate	80 Hz
N. channels	2
N. epochs	2000
Optimizer	SGD
Scheduler type	Reduce learning rate when a metric has stopped improving
Learning rate	0.004
Patience	25 epochs
Factor scale	0.2
Batch size	16, 32
Decay	0.001

Table 4.1: The hyperparameters used to train all EEGnet

To build the *Network Tree*, the classification of the unknown class for each network was tested with the following results:

- a rest epoch, if given as an input to $EEGnet_{LR}$, was classified as right;
- a right epoch, if given as an input to $EEGnet_{0L}$, was classified as rest;
- a left epoch, if given as an input to $EEGnet_{0R}$, was classified as right.

Each network started with a CSP transformation, and each CSP transformation was different for each network. Based on the results shown in the Tab. 4.2, the system behaves as follows: a sample was given as input in $EEGnet_{LR}$: if it was classified as left, the original epoch was given to $EEGnet_{0L}$ for the final classification in the rest or left movement; if it was classified as right, then the original epoch was given as input in parallel to two networks: $EEGnet_{0L}$ and $EEGnet_{0R}$. Based on the results, the truth table in Fig. 4.4 was applied to classify the epoch.

The network tree architecture is shown in Fig. 4.4. Tab. 4.3 shown the classification results using the *Network Tree*.

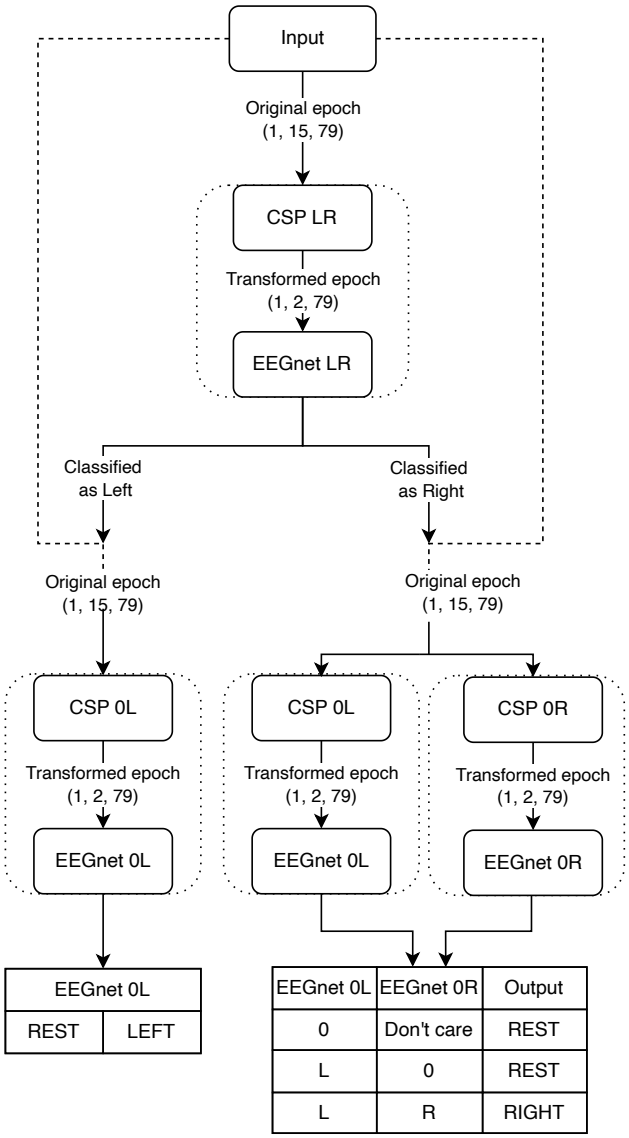


Figure 4.4: *Network Tree* - the first classification of the signal is performed by the $EEGnet_{LR}$ and then a classification check is carried out. If L is identified, the input signal is sent to the $EEGnet_{0L}$ for final classification between 0 or L. In case R is classified, the input signal is sent both to the $EEGnet_{0L}$ and $EEGnet_{0R}$. The result of their combined classification returns the result of the final classification between 0 or R. The input undergoes, before being sent to a network, a transformation by the associated CSP.

Network	Class	Support	Precision	Recall	F1	Accuracy	Unknown Class	Prediction
$EEGnet_{0LR}$	rest	574	0.62	0.25	0.24	0.43	-	-
	left	594	0.39	0.96	0.55		-	-
	right	606	0.72	0.28	0.29		-	-
$EEGnet_{LR}$	left	600	0.74	0.73	0.74	0.73	rest	0.46
	right	582	0.73	0.73	0.73			0.54
$EEGnet_{0L}$	rest	576	0.66	0.64	0.65	0.67	right	0.73
	left	609	0.67	0.69	0.68			0.27
$EEGnet_{0R}$	rest	568	0.67	0.46	0.55	0.63	left	0.42
	right	614	0.61	0.79	0.69			0.58

Table 4.2: Prediction metrics table for all networks used in *Network Tree*

Network	Class	Support	Precision	Recall	F1	Accuracy
Network Tree	rest	591	0.48	0.45	0.46	0.55
	left	586	0.57	0.61	0.59	
	right	596	0.58	0.59	0.58	

Table 4.3: Results of *Network Tree*

4.1.2 Results

For evaluating the model, Accuracy, Precision, Recall, and F1 metrics were used. In Tab. 4.2 the results are shown for all networks trained individually. In the last three rows of Tab. 4.2, the results for all two-class networks showed that the binary classification was more accurate related to the three-class network (last row). The $EEGnet_{RL}$, showed a good discriminability power between left and right movements, while when a rest signal was given as an input to this net, it created a light unbalancing classification result towards the right class. It was interesting to note that a right sample was classified as rest if it was given as input to $EEGnet_{0L}$, while a left sample was classified as right if was given as input to $EEGnet_{0R}$. Starting from these results, it was built the *Network Tree* combining all 2-classes network in a classification cascade. The proposed network achieved an overall accuracy of 0.55, greater than the chance [145] and improved performance with respect to $EEGnet_{0LR}$. Moreover, in $EEGnet_{0LR}$ the single classes were frequently misclassified, especially rest and right, and it was shown by looking at the recall metric. It should be emphasized that in the training of $EEGnet_{0LR}$, three classes were extracted with CSP instead of two, so the resulting multivariate signal consisted of 3 channels. For comparison, the results for the *Network Tree* are reported in the Tab. 4.3, showing a more stable behaviour of the network in terms of classification. This aspect is well represented by the F1 score, which takes into account both precision and recall.

This section aimed to present a new method for a motor-based BCI. In an online classification, the EEG signal is recorded while performing a movement of the L o R hand or while having a rest. In this field, there is a literature gap on the non-movement classification, also known as rest, but this is fundamental in online application since there are many non-movement periods while using a BCI [182]. The novelty lies in the concatenation of a network tree of able of correctly classifying not only left- and right-hand movements, but also the rest signals. Future work should test the system in real-time by including a pre-processing pipeline capable of quickly cleaning the signals [83, 85]. Furthermore, other tasks, such as motor imagery in emotions [2], need to be tested using different types of feature extracted by the EEG data, such as connectivity [4], independent components [184], or other features [185]. Moreover, future improvement of the *Network Tree* will include both subject-dependent and subject-independent cross-validation analysis, splitting subjects into training and test groups prior to training.

4.2 Deep Learning Architecture analysis for EEG-Based BCI

The literature on MI decoding and classification is full of different architectures [152], and many reviews compare the results of these researches [186, 187]. In [155], a hybrid model combining CNN and transformer architectures, ConTraNet, was constructed to increase classification performance in several EEG applications, including rest categorization. In [156], the VS-LSTM model was used to categorize limb MI using EEG signals, with a focus on discriminating between motor and resting states. In [158], the authors performed the classification on MI signals, open and closed eyes, using LDA, NB, and SVM, attaining an accuracy of 91.18 %, 95.41 %, and 99.51 %, respectively. Furthermore, in [159], the authors classified 2, 3 and 4-classes using EEGnet on MI and rest task, attaining an overall accuracy of 82.43 %, 75.07 %, and 65.07 % respectively. In [160], the authors classified MI signals and rest with an accuracy of 70.64 % across five classes. Few articles [188–190] tried to decode information under ME task and its decoding is primarily for telerehabilitation and for driving robotic arms in people who cannot move with hand-based sensors. The

purpose of this study is to understand, among the scientific literature, which is the best DL architecture designed for different domains and applied for ME classification. A comparison of three different types of DL architecture training the network from scratch, using the data from the PhysioNet and Upper Limb, preprocessed in the same way to make them similar, was provided. In this way, the comparison of results is not influenced by the bias introduced in each step of feature cleaning, extraction, or selection. To the best of our knowledge, no other studies have conducted a similar analysis, making it impossible to compare this achievement with others previously published in the literature.

4.2.1 Materials and methods

This section aims to compare different DL architectures and different public dataset EEG signals under ME for real-time BCI [182]. Two datasets and three different architectures were used. The public dataset [152] used are the following: (1) PhysioNet [86] and (2) Upper Limb Graz Dataset [163] for ME and the three DL architectures used are: (1) EEGnet [176], (2) Deep4Net [128] and (3) EEGITnet [127]. These networks were chosen because they were already implemented in the Braindecode library [191] and were suitable for these types of applications and data.

Dataset description Data from 105 individuals from the PhysioNet dataset [86] were used. The original dataset included 109 participants; however, four were eliminated due to varying trial counts and sampling rates. Subjects were recorded performing or imagining opening and closing their right or left hand for 4 seconds. There was a 4-second rest period before each exercise or imagined task. Each subject was recorded three times, with roughly seven motions in each category and one rest period between each. Data from all individuals from Upper Limb [163] were used. Subjects were recorded performing movement or rest for 3 seconds. Subjects were recorded for 10 runs with 42 trials per run. Dataset presented 6 movement and rest classes. For the PhysioNet dataset there are no demographics information; whereas the Upper Limb Graz dataset consists of EEG data from 15 healthy subjects between 22 and 40 years (average: 27 $\pm$5), with 9 female and 6 male. All subjects were right-handed excepted from the first subject.

Dataset preparation In this section "open and close left or right fist", "rest", "hand close", "hand open" and "rest" were considered from the PhysioNet and Upper Limb datasets, respectively. For the PhysioNet movement class was selected only the signal from the right hand, thus the data between the two datasets remain heterogeneous, as the Upper Limb dataset contains data from right-handed subjects. The training phase was performed in two classes, the "rest" and the "open/close" right hand. For the Upper Limb dataset, the training phase was performed on two classes, "rest"-"hand open/close". The class "open/close" was realized by combining the single classes "hand close" and "hand open" of the dataset for heterogeneity between the two datasets.

Preprocessing For each participant in the two datasets used, the same preprocessing pipeline was applied. Following some general rules about EEG preprocessing [83, 85, 140, 192], it was applied the following pipeline to preprocess the data in MATALB using EEGLAB [193]. A simple diagram of the preprocessing pipeline in the first part of Fig. 4.5 is shown. First, the location of the channels is loaded according to standard 10-20 [194], and the channels were selected based on the most recent finding about motor decoding [189], and with the well-known neuroscientific knowledge, using the following: FC1, FC3, FC5, FCz, FC2, FC4, FC6, C5, C3, C1, Cz, C2, C4, C6, CP5, CP3, CP1, CPz, CP2, CP4, CP6. The positions of those channels in Fig. 4.6 are shown.

Then, only EEG valid data was selected and the rest was excluded. After data exclusion, a Butterworth bandpass filter from 1 to 30 Hz was applied and subsequently the ICA algorithm was used to decompose the data into ICs. Finally, using the IC-label plug-in for EEGLAB [85], noisy components were excluded, according to the values reported in Tab. 4.4. Finally, using Python and the MNE library [177], the signal

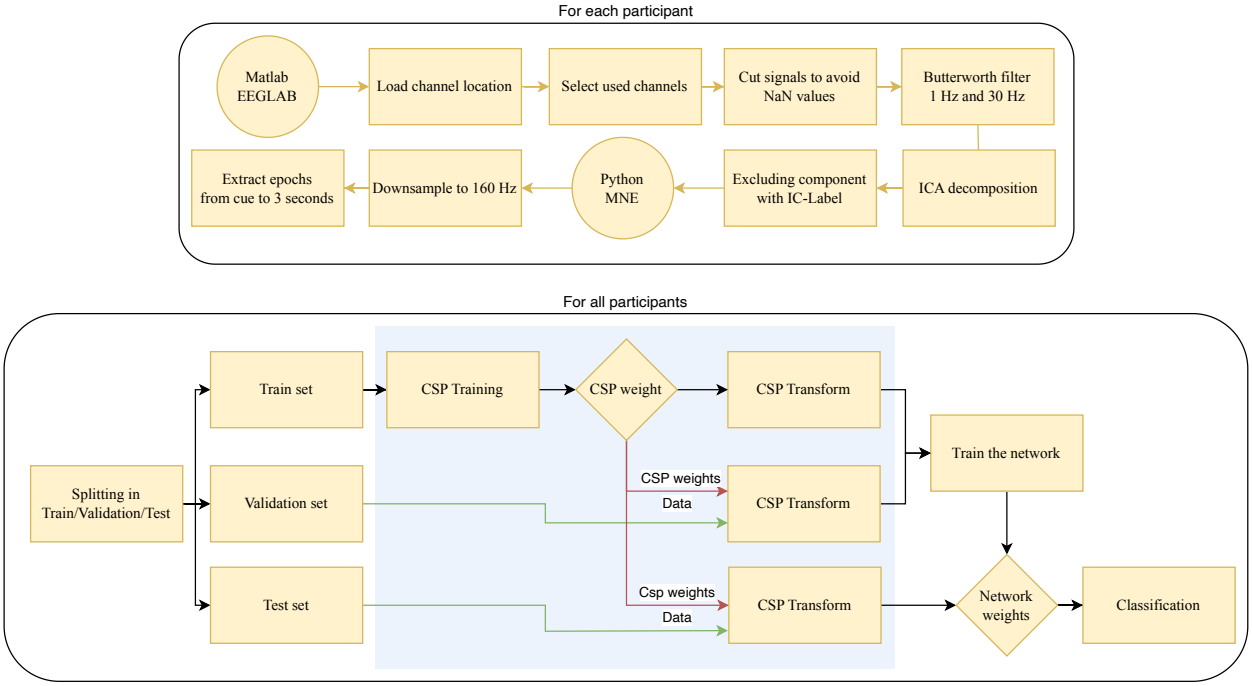


Figure 4.5: In the upper part, the pipeline applied to clean the signals is shown performed for each participant. The training and test phases are shown in the lower part after merging all samples in one set. The processes in the blue box are skipped if the CSP transformation is not applied.

Brain	Muscle	Eye	Heart	Line Noise	Channel Noise	Other
N	0.8	0.8	0.8	0.8	0.8	N

Table 4.4: Parameters for each IC-label category. The *"N"* letters mean that these components are not removed, while the other values represent the threshold to choose between leaving and removing the component if it was classified as the label.

was downsampled to 160 Hz to make the Upper Limb dataset close to PhysioNet, and finally the epochs were extracted based on movement events from the cue for 3 seconds of the signal. In this section the CSP algorithm [98] was applied to the data testing whether the classification results improved under the ME condition. CSP is a statistical method employed in signal processing and machine learning to improve the distinction between two classes by mapping the original data onto a new dimensional space. The technique operates by increasing the variance for one class and reducing it for the other, thereby creating effective spatial filters that accentuate the distinguishing characteristics between the classes. The CSP is used a lot in EEG signal classification tasks under MI conditions [195] and much less in ME tasks. The number of channels were set to 2 when CSP was applied; otherwise, it was set to 21. The data were used first without the CSP transform and then transformed with the CSP algorithm, to compare the results with this type of transformation, as shown in the Tab. 4.7 and 4.8. The reason to use CSP is to separate the classes before the learning process starts, allowing for the reduction of useless information, the weight of the architecture, and the speed of the learning process. In this section, it was used the threshold theorized with simulated data by [145] for the ME classification.

Deep learning models The classification phase was performed using the network already implemented in the Braindecode Python library [191]. The entire training and test phase process is shown in Fig. 4.5. Tab. 4.5 displays the training hyperparameters, which were chosen after numerous experimental trials and implemented across all networks. The SGD was utilized to enhance the convergence of the training, in conjunction with a dynamic LR that started at 0.004 and automatically decreased when there was no

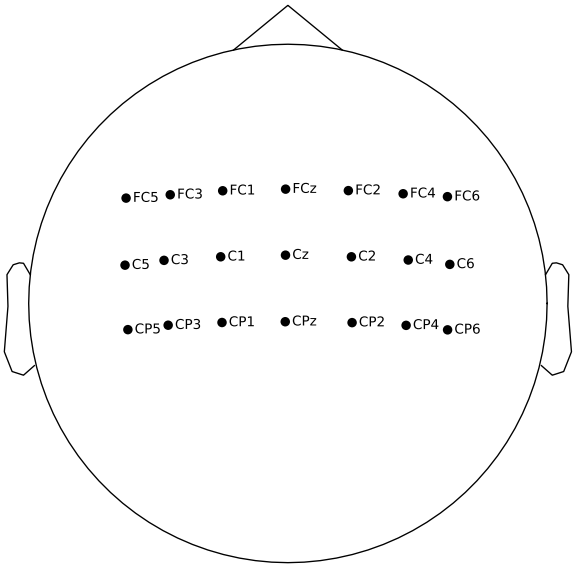


Figure 4.6: The channels used in this section. The channels are located in the retral part of the frontal area, on the central area and the anterior part of the parietal area, covering the motor areas. This process was done due to reduce the useless information not related to sensory-motor areas, improving the learning process.

Hyperparameter	Value
Sample shape	1, 21, 480
Sampling rate	160 Hz
N. channels	2 / 21
N. epochs	2000
Optimizer	SGD
Scheduler type	Reduce LR when a metric has stopped improving
Learning rate	0.004
Patience	25 epochs
Factor scale	0.2
Batch size	16, 32
Decay	0.001

Table 4.5: The hyperparameters used to train all Network

improvement after 25 inactive epochs. The reduction factor for the LR was established at 0.2. Furthermore, it was used a double batch size [181] that was adjusted when the network did not show any further enhancement even after the LR was reduced. Lastly, it was applied a decay rate of 0.001 to increase the network learning rate. This adjustment of the batch size during the learning process allows the network to enhance its performance.

4.2.2 Discussion and results

The curves for validation set accuracy, training set loss, and validation set loss from the two top-performing models in both datasets are displayed in Tab. 4.6. These curves exhibit a pattern that aligns with the scientific literature on the behaviour of accuracy and loss during learning [196]. A comparison of Tab. 4.7 and 4.8 reveal that among the used models, EEGnet has the best performance on both datasets. Furthermore, the EEGITnet model underperforms when compared to the EEGnet model, while the Deep4net has better results compared to EEGITnet but it has worse results concerning EEGnet. An analysis of the transformed CSP data indicates that all tested networks perform better without the CSP-transformed data. CSP is often

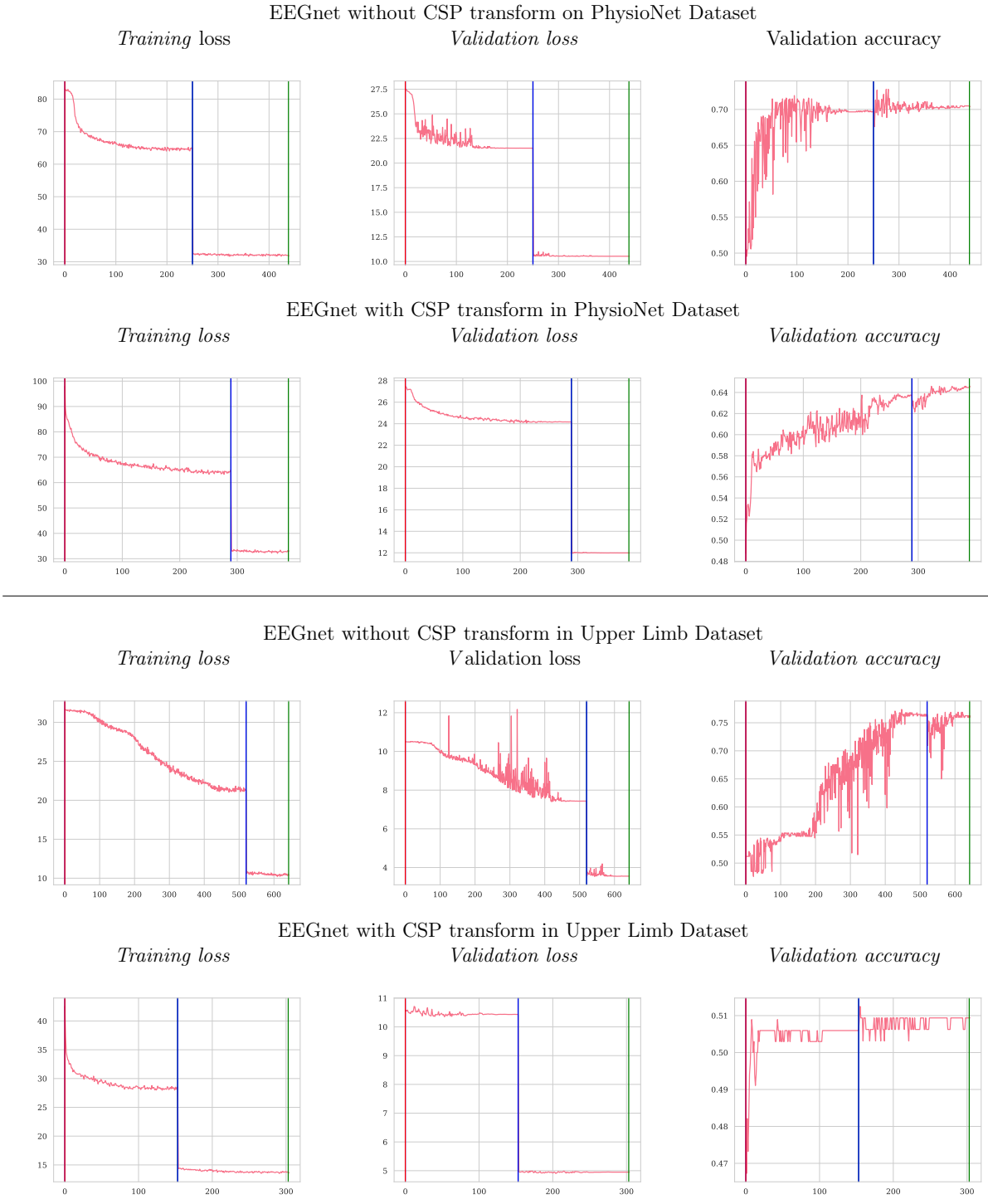


Table 4.6: The validation accuracy and the training and validation loss curves are plotted for the EEGnet. The best results were achieved from the EEGnet trained without CSP transformation, for both datasets. The blue line indicates the change in batch size during the learning process from 16 to 32, and the green line indicates when the early stopping function occurs. The first two rows illustrate the curve related to the PhysioNet dataset, whereas the subsequent two rows show the curve for the Upper Limb dataset, without and with CSP. Each column represents Training Loss, Validation Loss, and Validation Accuracy respectively for each training session.

Dataset	Network	Class	Support	Precision	Recall	F1	Accuracy
PhysioNet	EEGnet	rest	589	0.66	0.81	0.73	**0.70**
		Open/Close	593	0.76	0.58	0.66	
PhysioNet	Deep4Net	rest	595	0.00	0.00	0.00	0.50
		Open/Close	587	0.50	1	0.66	
PhysioNet	EEGITnet	rest	551	0.47	0.74	0.57	0.49
		Open/Close	631	0.55	0.28	0.37	
Upper Limb	EEGnet	rest	237	0.80	0.76	0.78	**0.77**
		Open/Close	213	0.74	0.79	0.77	
Upper Limb	Deep4Net	rest	221	0.49	1	0.66	0.49
		Open/Close	229	0.00	0.00	0.00	
Upper Limb	EEGITnet	rest	213	0.47	0.65	0.54	0.48
		Open/Close	237	0.52	0.34	0.41	

Table 4.7: Prediction metrics table for all networks used in *PhysioNet* and *Upper Limb* dataset without CSP. In bold, the best accuracy for each dataset and model.

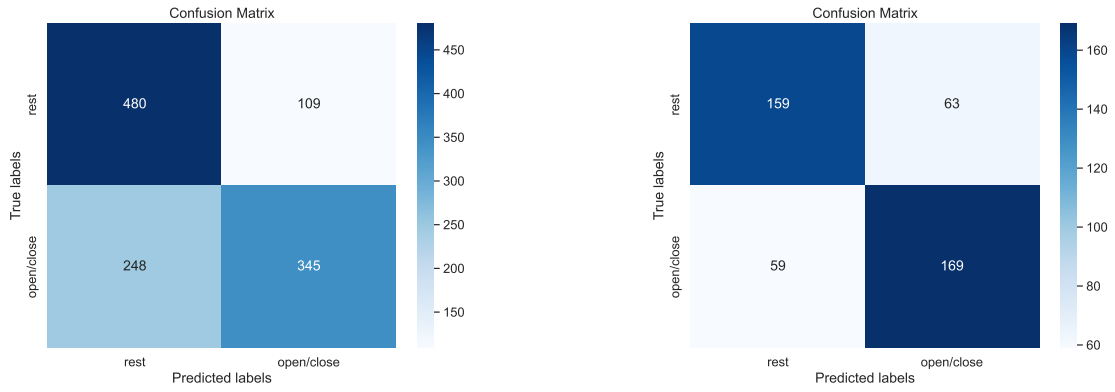
Dataset	Network	Class	Support	Precision	Recall	F1	Accuracy
PhysioNet	EEGnet	rest	572	0.74	0.43	0.54	**0.65**
		Open/Close	610	0.62	0.85	0.72	
PhysioNet	Deep4Net	rest	590	0.50	0.57	0.53	0.50
		Open/Close	592	0.50	0.44	0.47	
PhysioNet	EEGITnet	rest	602	0.54	0.34	0.42	0.49
		Open/Close	580	0.50	0.70	0.59	
Upper Limb	EEGnet	rest	240	0.55	1	0.71	**0.56**
		Open/Close	210	1.00	0.07	0.22	
Upper Limb	Deep4Net	rest	224	0.42	0.04	0.08	0.49
		Open/Close	226	0.50	0.94	0.65	
Upper Limb	EEGITnet	rest	244	0.53	0.25	0.34	0.48
		Open/Close	206	0.46	0.74	0.56	

Table 4.8: Prediction metrics table for all networks used in *PhysioNet* and *Upper Limb* dataset with CSP. In bold, the best accuracy for each dataset and model.

used to improve class separation in MI tasks [197], whereas there is no published work on the classification of ME tasks with CSP, suggesting that the CSP algorithm may not be suitable for ME tasks. Furthermore, the PhysioNet dataset contains a single class that encompasses both "open" and "close", while the Upper Limb dataset distinguishes between these two with separate classes. In this study for the Upper Limb Dataset, these two classes were combined into one to facilitate a comparison between the two datasets. The findings indicate that, given the same DL model and with CSP transformation, the Upper Limb dataset outperforms the PhysioNet dataset. This implies that the movements labeled as "open" and "close" are nearly indistinguishable from one another and could be considered as a single class relative to the "rest" condition. This also accounts for the performance of EEGnet on the PhysioNet dataset, which can effectively differentiate between the "open/close" and "rest" classes. For evaluating the model, Accuracy, Precision, Recall, and F1 metrics were used.

In Tab. 4.7 and Tab. 4.8 the results are shown for all networks trained individually and applied for a single dataset without CSP and with CSP, respectively.

This section aimed to evaluate three DL architectures, (1) EEGnet [176], (2) Deep4Net [128] and (3) EEGITnet [127], to classify the EEG signal under ME for real-time BCI. The network was trained from the start, using the same data on the same hardware to better compare the results with less bias introduced at each preprocessing step. According to the results of these tests, the best classification architecture for both datasets was EEGnet, which did not employ the CSP for feature extraction and instead sent the EEG input untransformed. The confusion matrix for the EEGnet trained on both dataset without CSP transformation is shown in Fig. 4.7 Future work should include an evaluation of the classification with EEGNet in the real-time domain and apply the same conceptualization in the Transformer architecture to find the best model to apply the presented method in a real-time BCI to control robotic hands for rehabilitation. The interpretation regarding a better result in the classification of the Upper Limb dataset is related in the acquisition: this dataset is much more recent and used a weight-discharging arm support, whereas the PhysioNet is a much older dataset that probably did not use any support in the rest sessions. Finally, future research will focus on interpreting the outcomes through a physiological lens, considering the brain regions engaged in movement. Future improvement will take into account other cognitive features such as emotions [2, 182] or P300 [198] and other types of signal transformation [4, 93, 199] to improve Human-Robot Interaction (HRI)/hci with brain signals together with CV [200].



(a) (b)

Figure 4.7: The confusion matrix of the EEGnet without CSP transform for both datasets, **(a)** the PhysioNet dataset and **(b)** the Upper Limb dataset.

4.3 The MOVING dataset

This section introduces MOVING, a Multi-Modal dataset of EEG signals and VG Hand Tracking. This dataset comprises neural EEG signals and kinematic data associated with three hand movements—open/close, finger tapping, and wrist rotation—along with a rest period. The dataset, obtained from 11 subjects using a 32-channel dry wireless EEG system, also includes synchronized kinematic data captured by a VG system equipped with two orthogonal LMCs. The use of these two devices allows for fast assembly (∼1 min), although introducing more noise than the gold standard devices for data acquisition. The study investigates which frequency bands in EEG signals are the most informative for motor task classification and the impact of baseline reduction on gesture recognition. Deep learning techniques, particularly EEGnet, are applied to analyze and classify movements based on the EEG data. This dataset aims to facilitate advances in BCI research and in the development of assistive devices for people with impaired hand mobility. This study contributes to the repository of EEG datasets, which is continuously increasing with data from other subjects, which is hoped to serve as benchmarks for new BCI approaches and applications.

The present work is focused on creating a multi-modal EEG dataset to classify the neural signals and to correlate neural information with three human-hand movements (open/close, finger tapping, and wrist rotation) modeled by a CV-based system named VG [1, 151, 201]. The conceptualization of this dataset follows the multi-modal approach of some previous works [203], but differentiated by using VG with two sensors for kinematic data acquisition at high precision. Furthermore, this article focuses on baseline analysis and the most informative frequency bands for motion classification using a deep learning technique operating on the PSD [7]. The use of these two devices, VG and Enobio® EEG dry, allow data acquisition to begin in ∼1 min, thus making the device effectively usable in an outdoor environment. However, dry electrodes introduce significantly higher noise than wet EEG devices [205]; the latter are considered the gold standard for EEG recording [206]. The acquisition modality, the acquired data, and the classification movement process using the DL architecture EEGnet [176] are detailed. This architecture has been developed to be useful in many BCI applications such as P300 cortical potentials movement-related and sensory-motor rhythms [207]. The protocol is carried out using a simple motor task, composed of sequences of hand movements tasks composed of MI action or ME. In the proposed work, according to the literature, three different movements (both MI and ME) are used: open/close, finger tapping, and wrist rotation. In addition, EEG data are recorded under rest conditions. The same movements performed in MI were followed by ME and rest for 10 min. The EEGnet, developed in the PyTorch environment, is trained and validated using this new dataset based on EEG signals in ME and rest conditions.

4.3.1 Methodology

This section aims to present a multi-modal dataset on MI/ME on kinematic and brain data acquired with VG and a wireless dry, fast-wearable EEG device (Enobio® EEG [3]). In addition, its goal is to analyze which frequency bands are more informative for ME classification and the role of the baseline reduction method. These devices were used in a semi-controlled environment to emulate real-life situations [208] where high noise and disturbances are present. The classification is performed with EEGnet [176] because recent studies have shown that this architecture is one of the best performing [7, 209]. Three classes of data were collected in both the ME and MI conditions: open/close, finger tapping, wrist rotation, and the rest condition using a dry EEG device and VG. These findings are made available as a starting point for the development of rehabilitation support systems for patients who have experienced hand injuries with loss of mobility. The EEGnet model acts on the collected EEG signals to recognize movements, even if not executed correctly, allowing the corresponding movements to reproduce on a robot arm. In this way, it allows the patient to have feedback for improving the motor skills.

[3] https://www.neuroelectrics.com/solutions/enobio/32

4.3.1.1 Data Structure

The dataset is composed of two main independent data structures: the EEG data and the VG data.

EEG Data The EEG signal is arranged as a 2D matrix, with rows representing channels (electrodes) and columns indicating time points (samples). The EEG recording involves 32 channels at 500 Hz throughout ∼10 min, thus generating a matrix ∼32 × 300,000. For the epoch creation, it is necessary to fix some markers to identify the occurring events. The markers are placed through synchronization between Unity `https://unity.com/`, which communicates with the VG, and the EEG software, NIC2 that saves the EEG signal. In general, events are a matrix in which each row represents an event associated with the timestamp and the label of the event. This configuration, along with metadata such as sampling rate, electrode positions, and event markers, enables in-depth analysis and understanding of brain activity. An example of the data collected with this structure is shown in Figure 4.8.

VG Data As explained in the introduction of this Chapter, the VG uses two LMCs, each generating a dynamic 4D numerical representation of the hand. The hand model is reproduced into frames, each frame representing a temporal snapshot of the hand at a certain point in time (Fig. 4.9). These frames collect detailed information on the recognized hands, including the precise positioning of 25 joints and their timestamp. Each joint represents the articulation between two bones or a terminal bone (fingertip) and provides precise coordinates in three dimensions (x, y, and z), as well as the calculation of velocity and acceleration. In addition to joint data, the system collects other hand-related information, such as the position and velocity of the palm. Furthermore, the VG was calibrated to a single reference system [150], and Figure 4.9 shows the basis vector of the VG. The former points from the palm to the fingers, whereas the latter indicates the palm's direction related to the controller. These vectors are used to calculate the inner product and create a rotation matrix for the hand. This matrix simplifies the calculation of the hand's orientation angles (yaw, pitch, and roll), as well as the angle between any two bones. Velocity and acceleration are also determined for each joint by combining neighboring frames. Finally, the two LMCs have been calibrated to ensure that the views of both LMCs are aligned, hence improving data integration, and the data provided by the optimal viewer are used based on roll angle. This algorithm ensures that the data from the optimal viewer are used all the time [150]. Moreover, to handle the VG data, a new Digital Imaging and COmmunications in Medicine (DICOM)-based Information Object Definition (IOD) was also defined and designed [210] for Quantitative Physical Rehabilitation (QPR) due to its data structure generated by a multiple-source hand-tracking device. DICOM is a standard used in medical informatics for the storage and transmission of medical images and related information, and this new integration for QPR allows data to be stored and shared in a common format in the scientific and medical community for hand traking and, in generale, multiple-source device.

4.3.1.2 Data Acquisition

The protocol used in this section is similar to [154], in which the participant was seated in a comfortable chair, in front of a 20″ screen following the instructions visualized on the monitor. The participant wears an Enobio® EEG system with 32 dry electrodes (P7, P4, Cz, Pz, P3, P8, O1, O2, T8, F8, C4, F4, Fp2, Fz, C3, F3, Fp1, T7, F7, Oz, PO4, FC6, FC2, AF4, CP6, CP2, CP1, CP5, FC1, FC5, AF3, PO3), while the VG device is placed under the right arm to record kinematic data. Figure 4.10 (a) depicts the experimental procedure and Figure 4.10 (b) shows the acquisition environment.

Eleven healthy right-handed subjects (ten men and one woman; average age 28.27 ± 9.94) were recorded for 10 min. Before the recording, the consent information was signed by each of them. The protocol consists of a sequence of tasks, each composed of a series of actions shown on the screen to guide the participant in the movements (ME, MI, and rest). In Figure 4.11, the images depict the movements used in this section. Each action starts with a fixation cross displayed for two seconds, followed by an instruction to perform either

imagined or executed movement shown for the subsequent six seconds. Each task is composed of a triplet of actions, with movements always presented in the same sequence: rest, MI, and ME. The protocol consisted of a triplet of tasks, the first involving imagining and performing the open/close movement, followed by the imagined and executed wrist rotation, and finally the imagined and executed finger tapping. In Fig. 4.12, the structure of the protocol is reported, and in Figure 4.13, the instructions are shown with one repetition of the protocol. The participants were shown the instructions on the screen and asked to continue performing the movement (MI or ME) until the next instruction appeared. Each triplet of all types of movement was repeated eight times, corresponding to 10 min.

4.3.1.3 Preprocessing Data

The data were processed using the MNE [177] library for Python, and were divided into two parts: general preprocessing and specific preprocessing, as shown in Figure 4.14. The preprocessing pipeline started loading the channel location following the 10-20 international system. The first step of the general preprocessing included a Common Average Referencing (CAR) and a band-pass Butterworth filter between 1 Hz and 100 Hz. After that, the ICA weights were computed using the Extended Infomax algorithm [211] with 500 iterations, and with IC-Label [85], the "eye" components with probability greater than 0.9 were removed. Subsequently, artifact removal was applied [212] ("muscle" components were removed with a probability greater than 0.9). The average of components for each subject identified as "eye" or "muscle" component removed are 0.2 and 1.3, respectively. Finally, the signal was downsampled at 256 Hz. For this study, neither MI classes nor ME of finger tapping movement were analyzed in depth.

In the second part of the preprocessing, defining a specific preprocessing pipeline, a different band-pass filter was applied each time, and on the resulting signal, the baseline reduction method (subtraction of the EEG signal during fixation cross) was either applied or not. Furthermore, the epochs of the ME open/close, wrist rotation, and rest were extracted and combined in pairs. Moreover, in a further test, the two classes open/close and wrist rotation were combined to create a single class, "movement", to see if it was more distinguishable from the rest condition. The result of these combinations (two baseline conditions, six bandpass filters, three movement classes, and one rest) 4.10, corresponds to 46 trials. Then, the epochs aligned with the signal cue are segmented into pieces of one second, starting from the 2nd second for the open/close and wrist rotation epochs and the 3rd second for the rest epochs. The reason for this drop for the rest cue is to avoid interference with previous movements and to give time to start movements (open/close and wrist rotation). It was decided to exclude the 1st second for the movement because the movement did not start exactly with the appearance of the trigger but took a variable amount of time depending on the individual, as can be seen from Figures 4.15–4.18. Furthermore, the participants reported they did not perform the first movement correctly. For the "rest" task, again based on the participants' feedback and the analysis of the kinematic data, the first two seconds were excluded as the movement was not stopped instantaneously while the cross appeared but lasted for a few times even during "rest". This new type of multi-modal dataset not only allows the extraction of possible patterns between the two modes, but also the preprocessing of data from one mode based on information extracted from another. In this specific case, the cut-off times of the EEG signal were chosen based on the kinematic data of the VG, without performing a combined analysis. Thus, the kinematic data made it possible to avoid biases due to the recording protocol, which could not have been avoided in the absence of kinematic data. During the epochs extraction after cutting the signals, a data augmentation process was applied extracting one second of signal shifting only by 0.50 s from the following epochs, overlapping each epoch with the following of 50 %. Then, a train/validation/test split was performed to create the three sets for the neural network training, allocating 60 %, 20 %, and 20 % of the original data, respectively. Finally, the PSD from each sample was computed using Welch's Method [99] and used for training the network. Figure 4.14 shows the corresponding pipeline.

The data from the VG are stored in a JSON file. The LMCs tracks hand movements in a 3D coordinate system. Using this raw data, the LMCs software, (Ultraleap Gemini v5.20.0) constructs a numerical model

of the hand represented as a stick figure that mirrors the anatomy of the human hand, as illustrated in Figure 4.9. Specifically, the LMCs data are structured into frames [150]. Each frame is a unique moment in time, and includes all of the captured data by the controller at that instance. Within each frame, details regarding detected hands are included. This encompasses information such as hand identifier, palm location (x, y, z coordinates), palm speed, palm orientation (the direction the palm is oriented towards), and palm trajectory (the path the palm is moving along). For the hand, comprehensive data regarding each finger is supplied. This consists of details like the finger's unique identifier, type of finger (thumb, index finger, middle finger, ring finger, pinky finger), the position of the fingertip (x, y, z coordinates), finger's direction, and velocity. Figures 4.15 and 4.16 show the horizontal and vertical LMCs trajectory, respectively, and Figures 4.17 and 4.18 show the horizontal and vertical LMCs velocity.

In this section, the EEGnet was chosen over other architectures, for its performance in classifying motion in public datasets [7].

4.3.2 Results

This section summarizes the analyses of the data acquired according to the procedures in the previous sections. Several frequency bands were analyzed, and the entire band [1–45 Hz] was filtered in the preprocessing phase to separate the sub-bands delta [1–4 Hz], theta [4–8 Hz], alpha [8–12 Hz], beta [12–30 Hz], and gamma [30–45 Hz]. Tests also include analyses with and without baseline removal. The classification step was accomplished using the EEGnet implemented in the Braindecode Python library [174]. The block training process in Figure 4.14 shows the full training and testing process.

Table 4.9 illustrates the training hyperparameters, whose values were selected following extensive experimental trials. The SGD was used to increase training convergence, along with a dynamic LR that began at 0.004 and automatically dropped when there was no improvement after 25 idle epochs. The reduction factor for the LR was set at 0.2. Furthermore, it was employed a double-value batch size [181], which was modified when the network showed no further enhancement even after the LR was reduced. Finally, it was used a decay rate of 0.001 to boost the network learning rate. The change in the batch size throughout the learning phase enabled the network to improve its performance. For model evaluation, Accuracy, Precision, Recall, and F1 metrics were used. They are based on TP, TN, FP, and FN. Concerning rest vs. movement, the best result was obtained by analyzing the whole frequency range [1–45 Hz] with baseline reduction, achieving an accuracy of 0.64, whereas when analyzing rest vs. single movements, the frequency band that allows us to achieve better results was the theta band. The baseline reduction allowed us to achieve an accuracy of 0.61 and 0.55 for rest vs. open/close and open/close vs. wrist rotation, respectively. For rest vs. wrist rotation, the baseline reduction was not as effective as in the previous tests, obtaining an accuracy value of 0.58, above the chance value [145], but three percentage points less than the best result, shown in Table 4.10. To verify the robustness of the results obtained in the different combinations of tasks that exceeded the threshold [145], a five-group k-fold cross-validation was performed and the results are shown in Table 4.11. The stability in the average value and standard deviation confirms the robustness of the trained models. For the kinematic data acquired by the VG, the trajectories and velocities of each fingertip on each axis and for each of the two LMCs, horizontal and vertical, were calculated and reported in Figures 4.15–4.18. From the graph of the z-component trajectories and velocities (the Cartesian reference system is oriented as in Figure 4.9), it can be seen that the more effective movements along this axis are open/close and finger tapping; while the wrist rotation movement is mostly affecting the x and y components. Observing the open/close z-component in the chosen time window, the individual performs on average 3 open/close movements. Regarding wrist rotation, the number of times a complete movement is performed is 3, as it is observed mostly on the y-axis. Regarding finger tapping, the most affected components are the y and z. All fingers are involved in finger tapping but the greatest contribution is by the index finger, mostly observable from the z-component. In the initial rest phase, the hand did not stop immediately, but took a variable time before arresting, depending on the individual and $\sim$1–2 s to maintain a constant position. Finally, for a preliminary visual analysis,

Figure 4.19 reports the topoplots indicating the PSD distribution of one of the subjects in the dataset for each of the movements reported along the timeline. The analysis of the topoplots shows that the activation in certain areas is more pronounced in the ME task than in the MI task, and almost absent in the "rest" task.

Hyperparameter	Value
Sample shape	1, 32, 129
Sampling rate	256 Hz
N. channels	32
N. epochs of training	2000
Optimizer	SGD
Scheduler type	Reduce LR when a metric has stopped improving
Learning rate	0.004
Patience	25 epochs
Factor scale	0.2
Batch size	16, 32
Decay	0.001
Baseline reduction	Yes/No

Table 4.9: The EEGnet hyperparameters.

4.3.3 Discussion

In this section, EEG signals and kinematic data related to MI and ME of open/close, wrist rotation and rest conditions were collected by 11 healthy subjects, stored in the MOVING dataset, and analyzed (those related to ME). Moreover, a new class called "movements" was built grouping open/close and wrist rotation, as shown in Figure 4.11. In the first step of the analysis, "rest" was chosen as the class of comparison between the two selected movements (merged in one class), in view of its usage with robot-assisted BCI. The tests carried out focused on finding the correct classification of the above-mentioned classes. Therefore, the following experiments were performed: rest vs. open/close, rest vs. wrist rotation, and open/close vs. wrist rotation. Despite the high-level noise affecting data (see Figure 4.8), using the preprocessing pipeline (Figure 4.14) and EEGnet in the frequency domain, the results are above the literature threshold [145], as reported in Table 4.10, except in one case (open/close vs. wrist rotation). These results are in line with previous recent studies [7]. It was also sought to determine which frequency band was more informative using PSD and found it was the theta band. Furthermore, the plots in Figures 4.15–4.18 show the kinematics of the gestures used in this dataset for one of the analyzed subjects. By observing the trajectory and velocity, it is possible to reconstruct the movements of all the fingertips. The synchronism between the EEG data and the kinematic data also allows us to focus on part of the EEG signal with more information about the movement, as can be observed from the trajectories and velocities plots. It can also be noticed that the rest in the first few seconds is influenced by movements and, in the same way, the movement epochs in the first few seconds are in the rest position. In fact, before starting the execution, a subject-dependent delay, due to the processing of the stimulus, was present. The MOVING dataset, to the best of our knowledge, is the first dataset that contains a numerical model of the moving hand and the corresponding EEG motion-related signals. It aims to become the benchmark in the scientific community for multi-modal analysis of multi-modal signals produced by movement. Finally, from Figure 4.19, it can be observed that there is a gradual decrease in PSD in the transition from ME to MI and from MI to "rest". This figure correlates with the cognitive effort required for the three tasks. One can see how the PSD between ME and MI presents the same spatial pattern, differing in the value of the PSD, as represented in the legend on the right. Moreover,

Table 4.10: Overview of the results. The movement class represents the merging of open/close and wrist rotation epochs. The best results for each combination of *baseline* and *filter* for each binary classification are in the box. In orange are the results that go above the chance threshold [145]. The box around the value indicates the best results for the classification.

Baseline	*Filter*	*Classes*	Precision	Recall	F1 Score	Support	Accuracy	*Classes*	Precision	Recall	F1 Score	Support	Accuracy
Yes	*1–45*	*rest*	0.71	0.51	0.59	346	0.64	*rest*	0.61	0.56	0.59	177	0.59
		movement	0.60	0.78	0.68	325		*open/close*	0.57	0.61	0.59	165	
No	*1–45*	*rest*	0.60	0.62	0.61	341	0.60	*rest*	0.54	0.73	0.62	165	0.57
		movement	0.60	0.58	0.59	334		*open/close*	0.64	0.43	0.52	182	
Yes	*1–4*	*rest*	0.54	0.7	0.61	346	0.54	*rest*	0.52	0.68	0.59	177	0.51
		movement	0.53	0.37	0.44	325		*open/close*	0.49	0.33	0.39	165	
No	*1–4*	*rest*	0.54	0.64	0.59	341	0.55	*rest*	0.50	0.71	0.58	165	0.52
		movement	0.55	0.45	0.50	334		*open/close*	0.57	0.35	0.44	182	
Yes	*4–8*	*rest*	0.60	0.57	0.58	389	0.57	*rest*	0.64	0.55	0.59	190	0.61
		movement	0.55	0.58	0.57	358		*open/close*	0.58	0.67	0.62	179	
No	*4–8*	*rest*	0.58	0.56	0.57	382	0.58	*rest*	0.51	0.63	0.57	183	0.53
		movement	0.57	0.59	0.58	379		*open/close*	0.56	0.44	0.49	195	
Yes	*8–12*	*rest*	0.56	0.51	0.53	403	0.54	*rest*	0.53	0.68	0.60	191	0.55
		movement	0.52	0.56	0.54	376		*open/close*	0.59	0.43	0.50	201	
No	*8–12*	*rest*	0.57	0.48	0.52	414	0.54	*rest*	0.53	0.51	0.52	191	0.54
		movement	0.51	0.60	0.55	369		*open/close*	0.55	0.57	0.56	201	
Yes	*12–30*	*rest*	0.58	0.60	0.59	405	0.57	*rest*	0.54	0.71	0.61	198	0.55
		movement	0.56	0.55	0.56	383		*open/close*	0.57	0.39	0.46	198	
No	*12–30*	*rest*	0.58	0.52	0.55	405	0.56	*rest*	0.52	0.71	0.60	198	0.53
		movement	0.55	0.61	0.57	383		*open/close*	0.55	0.35	0.43	198	
Yes	*30–45*	*rest*	0.50	0.32	0.39	405	0.48	*rest*	0.48	0.55	0.51	198	0.47
		movement	0.48	0.66	0.55	383		*open/close*	0.47	0.39	0.43	198	
No	*30–45*	*rest*	0.52	0.44	0.47	405	0.50	*rest*	0.48	0.62	0.54	198	0.48
		movement	0.49	0.57	0.52	383		*open/close*	0.47	0.34	0.39	198	
Yes	*1–45*	*rest*	0.65	0.49	0.56	172	0.60	*open/close*	0.52	0.58	0.55	128	0.53
		wrist rotation	0.56	0.71	0.63	157		*wrist rotation*	0.53	0.48	0.50	128	
No	*1–45*	*rest*	0.57	0.53	0.55	172	0.54	*open/close*	0.50	0.43	0.46	128	0.50
		wrist rotation	0.52	0.56	0.54	157		*wrist rotation*	0.50	0.58	0.54	128	
Yes	*1–4*	*rest*	0.58	0.63	0.61	172	0.57	*open/close*	0.56	0.45	0.50	128	0.55
		wrist rotation	0.55	0.50	0.52	157		*wrist rotation*	0.54	0.65	0.59	128	
No	*1–4*	*rest*	0.58	0.60	0.59	172	0.56	*open/close*	0.49	0.46	0.48	128	0.49
		wrist rotation	0.54	0.52	0.53	157		*wrist rotation*	0.49	0.52	0.51	128	
Yes	*4–8*	*rest*	0.56	0.64	0.60	183	0.58	*open/close*	0.56	0.67	0.61	153	0.55
		wrist rotation	0.61	0.53	0.57	195		*wrist rotation*	0.54	0.43	0.48	141	
No	*4–8*	*rest*	0.57	0.72	0.63	180	0.61	*open/close*	0.55	0.38	0.45	159	0.51
		wrist rotation	0.67	0.51	0.58	203		*wrist rotation*	0.48	0.65	0.55	139	
Yes	*8–12*	*rest*	0.52	0.51	0.52	194	0.52	*open/close*	0.53	0.61	0.56	157	0.51
		wrist rotation	0.52	0.53	0.52	193		*wrist rotation*	0.49	0.41	0.45	144	
No	*8–12*	*rest*	0.52	0.50	0.51	191	0.53	*open/close*	0.52	0.24	0.33	164	0.47
		wrist rotation	0.54	0.56	0.55	201		*wrist rotation*	0.46	0.74	0.56	141	
Yes	*12–30*	*rest*	0.54	0.55	0.55	191	0.55	*open/close*	0.57	0.49	0.53	164	0.52
		wrist rotation	0.57	0.56	0.56	201		*wrist rotation*	0.49	0.56	0.52	141	
No	*12–30*	*rest*	0.53	0.48	0.50	191	0.54	*open/close*	0.55	0.40	0.46	164	0.50
		wrist rotation	0.55	0.59	0.57	201		*wrist rotation*	0.47	0.62	0.53	141	
Yes	*30–45*	*rest*	0.45	0.42	0.43	191	0.47	*open/close*	0.57	0.45	0.50	164	0.52
		wrist rotation	0.48	0.51	0.50	201		*wrist rotation*	0.49	0.60	0.54	141	
No	*30–45*	*rest*	0.46	0.52	0.49	191	0.47	*open/close*	0.55	0.42	0.48	164	0.50
		wrist rotation	0.48	0.42	0.45	201		*wrist rotation*	0.47	0.60	0.53	141	

the "finger tapping" ME does not show any movement-related pattern (Figure 4.19), and this result explains the difficulty of EEGnet in classifying this task compared to "rest", which on the contrary can be easily performed with the other two ME tasks. The examination of the PSD spatial distribution across various ME, MI, and "rest" conditions is conducted on a single subject to illustrate the potential analysis that can be carried out. This calculation excludes the first second of the ME/MI condition and the initial two seconds of the "rest" condition: this exclusion is determined based on kinematic data analysis and on feedback reported by subjects, which indicated when subjects began and finished movements. The MI and "finger tapping" tasks were not analyzed with EEGnet, but only used for representing their topoplots shown in Figure 4.19. Here, it was provided additional data and tables that support the findings of this study.

This study aimed to introduce MOVING, a multi-modal dataset in a semi-controlled environment, consisting of EEG signals and kinematic data collected using VG during ME, MI, and rest conditions. The EEG data contained a high noise level due to dry electrodes: dry electrodes make the montage faster than wet electrodes. Furthermore, the influence of baseline reduction and frequency band analysis on ME classification was examined for only two ME movements and the rest condition, using PSD data extracted with Welch's method. The results, reported in Table 4.10, show that, for rest vs. movement (open/close and wrist rotation) classification, the most informative frequency band is [1–45 Hz], and the classification is supported

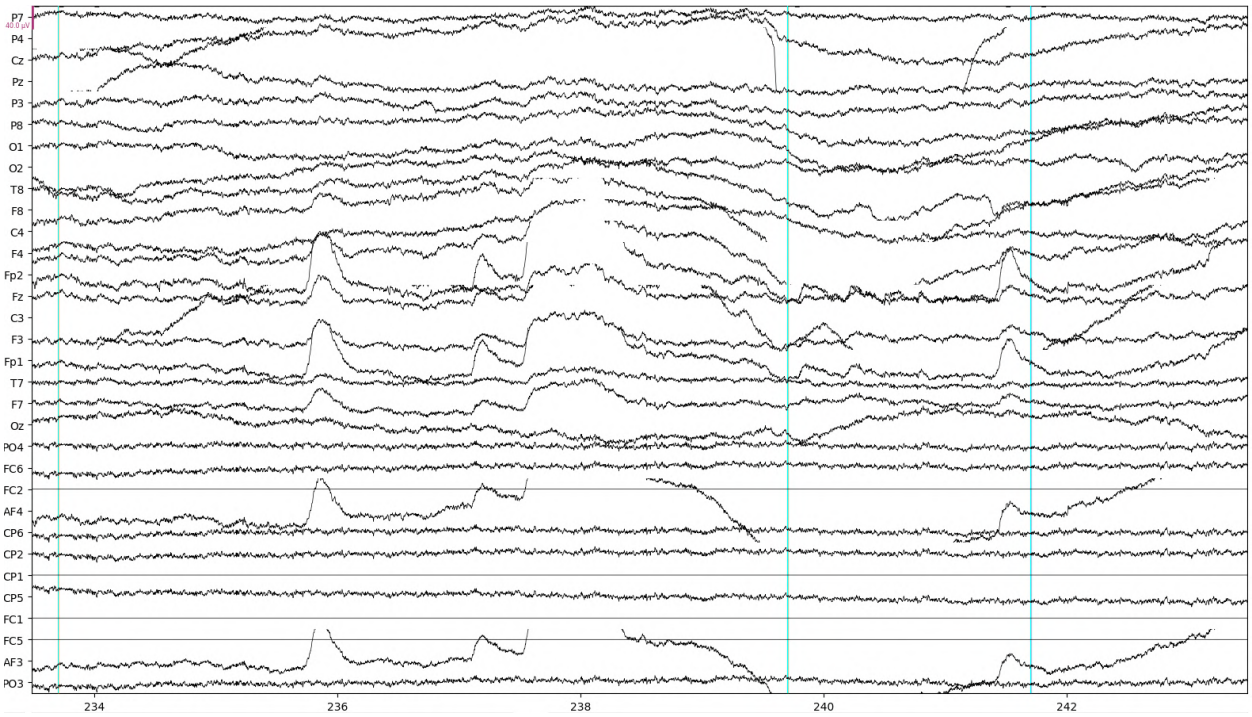
k-Fold	**rest/Mov**	**rest/Wrist**	**rest/Openclose**
1-fold	0.59	0.63	0.61
2-fold	0.62	0.61	0.60
3-fold	0.62	0.58	0.61
4-fold	0.60	0.59	0.60
5-fold	0.64	0.61	0.61
average ± dev.st	*0.61 ± 0.02*	*0.60 ± 0.02*	*0.61 ± 0.01*

Table 4.11: Results performing 5-fold cross-validation on the best models reported in Table 4.10.

if the data are corrected by baseline reduction. For single movement detection, theta was the most informative frequency band for both movements, but the network performed better with the baseline reduction for rest vs. open/close, and without baseline reduction for rest vs. wrist rotation. Finally, the possibility of using the same network for the open/close vs. wrist rotation classification was also investigated, but the results were not above the chance [145]. Moreover, these findings are in line with previous studies that used the same conditions [7]. The kinematic data collected with the VG were also analyzed. The plots shown in Fig. 4.15–4.18 suggest that: (1) movements were not stopped immediately with the appearance of the trigger; (2) movements were not started as the start trigger appeared, but needed some latency; and (3) movements were well reconstructed based on trajectories and velocities. On these considerations, the EEG signals were cut before the analysis, to reduce the bias due to the late start or stop of the movements. These considerations could be useful for the design of future experimental protocols to contain cleaner, bias-free, multi-modal datasets with simultaneous recording of EEG signals and kinematic data. The topoplots representing the scalp PSD in the 1–45 Hz frequency band over the analyzed time interval (Figure 4.19) show that the activation of certain areas during ME (in two over three movements recorded) is more pronounced than during MI or "rest". This initial analysis must be expanded by considering prior observations to conduct a comprehensive evaluation across all subjects. The MOVING dataset opens the possibility to investigate not only the preprocessing method and the characteristics of the hand movement extracted by the EEG signals, but also the relationship with the kinematic data acquired with the VG, here started with topoplot analysis. Future work will be dedicated to analyzing the MI part and building a portable system for safe operations in a work environment [5], improving the preprocessing steps by new fast algorithms [83, 214], and investigating new models to analyze EEG data [4, 215] to improve the performance classification in a real-life environment, such as to move a robotic arm in a dangerous workplace [216], or to interact with collaborative assistive devices [200, 217]. Finally, in the future, the MOVING dataset will be increased by including more subjects (men and women, left- and right-handed subjects) to make it more robust.

The studies reported in this chapter were related to the analysis of EEG signals acquired during ME. These experiments have contributed to an improved understanding of the way in which the EEG signals could be used to decode motor activities and have led to new techniques for the classification of such signals. The *network tree* architecture, composed of several concatenated EEGnets, was presented and shown to have an acceptable performance in classification of EEG signals associated with right-hand movement, left-hand movement, and the "resting state". These results show the improvement of the classification results of the EEG signals in these three classes. This approach, which divides the problem into several binary classes, could contribute to the creation of active and passive systems with better performance BCI. The second experiment reviewed the published preprocessing and classification pipeline related to ME with EEG signals. The results show the optimal combination for the classification of EEG ME through a comparative analysis of various DL architectures and two different preprocessing pipelines. The study showed that, on two well-known datasets, the best performance could be obtained using the data not transformed with CSP, and using the EEGnet architecture. The third study presented MOVING, a new multimodal dataset that combines EEG signals with kinematic data acquired by VG, while performing different movements ME

and MI. This section also presented the application of the same preprocessing and classification pipelines discussed in the previous work with highly noisy data acquired using a wearable wireless device EEG. The results show the models' ability to appropriately classify EEG signals from movement and rest periods even under less controlled, noisier conditions in signal acquisition. These studies have shown the importance of ME as an experimental paradigm for the development and validation of ME models in the analysis of EEG signals. The use of the paradigm of ME has several great advantages: (1) *Ground Truth Labels Verification*: Unlike other paradigms, such as MI o other pure cognitive tasks, in the ME condition, it is possible to have an objective and exact verification of the labels associated with the EEG signals, since the performed movements can be directly observed and accurately measured; (2) *Identification of the movement start, end, direction, and force*: for instance, kinematic data recording devices, such as the VG system used for MOVING dataset, it gives numerical hand truth identification of the start, end, direction, and force of every movement. This detailed information can be correlated with the characteristics of the EEG signals, thus allowing the identification of specific neural patterns associated with different aspects of the movement; (3) *correlation with EEG signal features*: kinematic data in conjunction with the EEG signals allow a deeper investigation of neural activity associated with movement characteristics. This correlation is essential to understand the neural mechanism of motor control and, in turn, would assist in the development of an effective DL model capable of decoding motor intention from EEG signals with improved accuracy. The experimental studies presented in this chapter show that ME is an effective experimental paradigm for the development of DL models for EEG signal analysis. The possibility of verifying labels, exactly identifying movement characteristics and relating them with EEG signals makes the ME paradigm a valuable tool for the development of BCI and neuroinformatics. Moreover, devices such as the VG will enable new perspectives on the development of multimodal datasets that include neural and kinematic information that may form a very solid basis for understanding and decoding movement-related brain activity.



(a)



(b)

Figure 4.8: EEG raw data before (**a**) and after (**b**) a cleaning band-pass [1–45 Hz] filtering. In (**a**), the high noise level makes it difficult to visualize high-amplitude brain signals. Vertical lines represent the triggers for rest, fixation cross, and open/close movement, respectively.

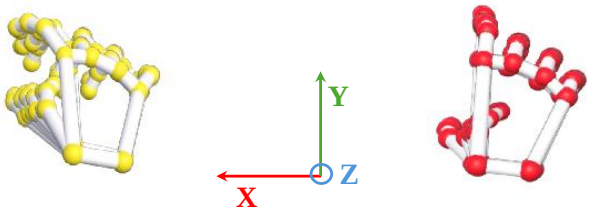


Figure 4.9: Model based on data collected by VG. The left part shows the vertical view; the right part shows the horizontal view. The Cartesian reference system is reported in the centre.



(a) (b)

Figure 4.10: Acquisition environment scheme (**a**) and real acquisition environment (**b**) used in this section.

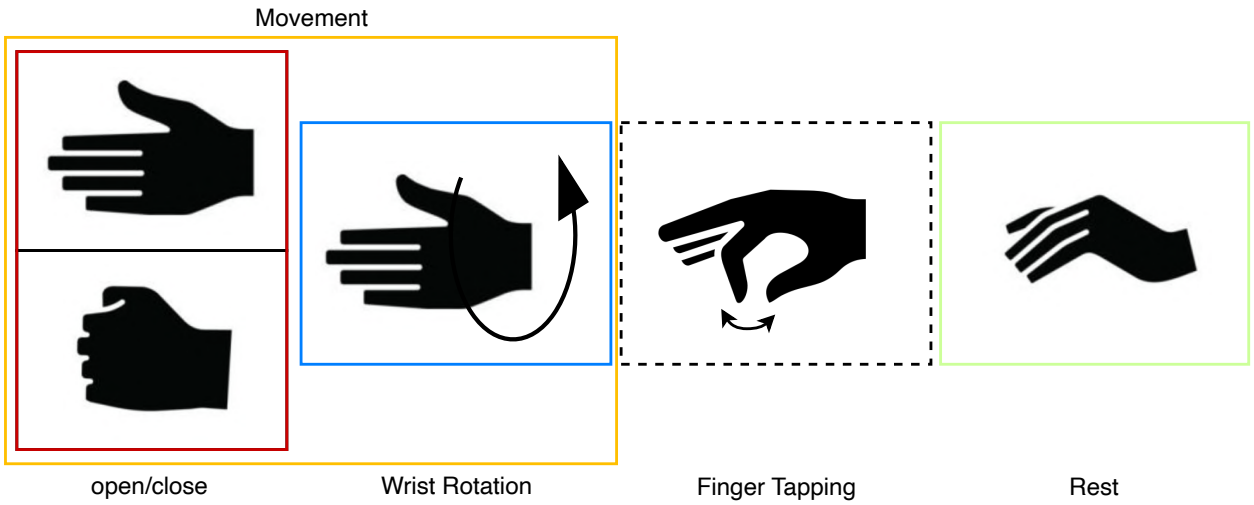


Figure 4.11: The analyzed movements. The dotted line represents the movement acquired but not analyzed in this section. The class "movement" is created by merging the open/close and wrist rotation classes.

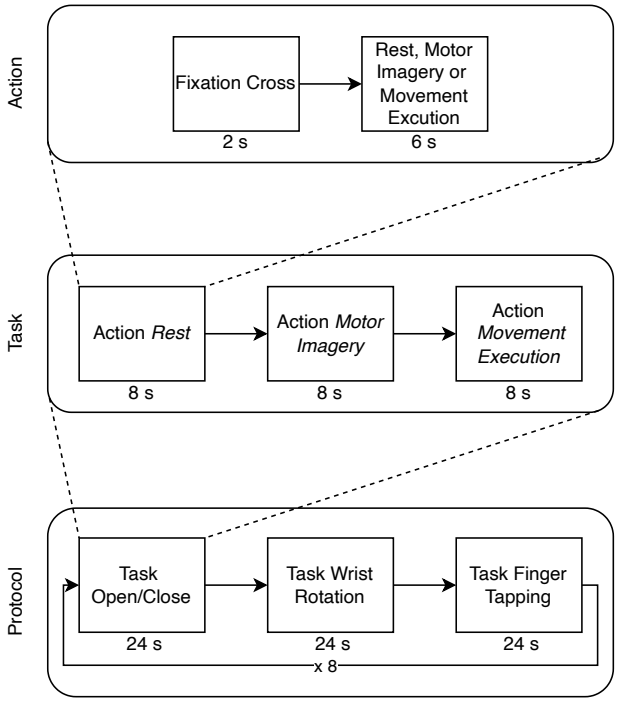


Figure 4.12: The used hand movement protocol.

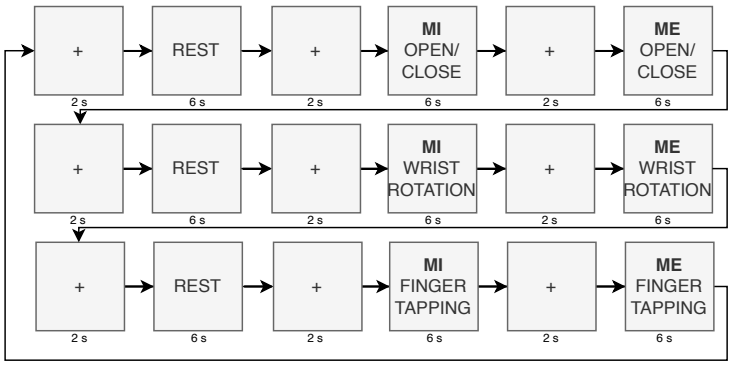


Figure 4.13: Instructions shown to participants for both MI and ME. The flow stops when 8 repetitions are reached.

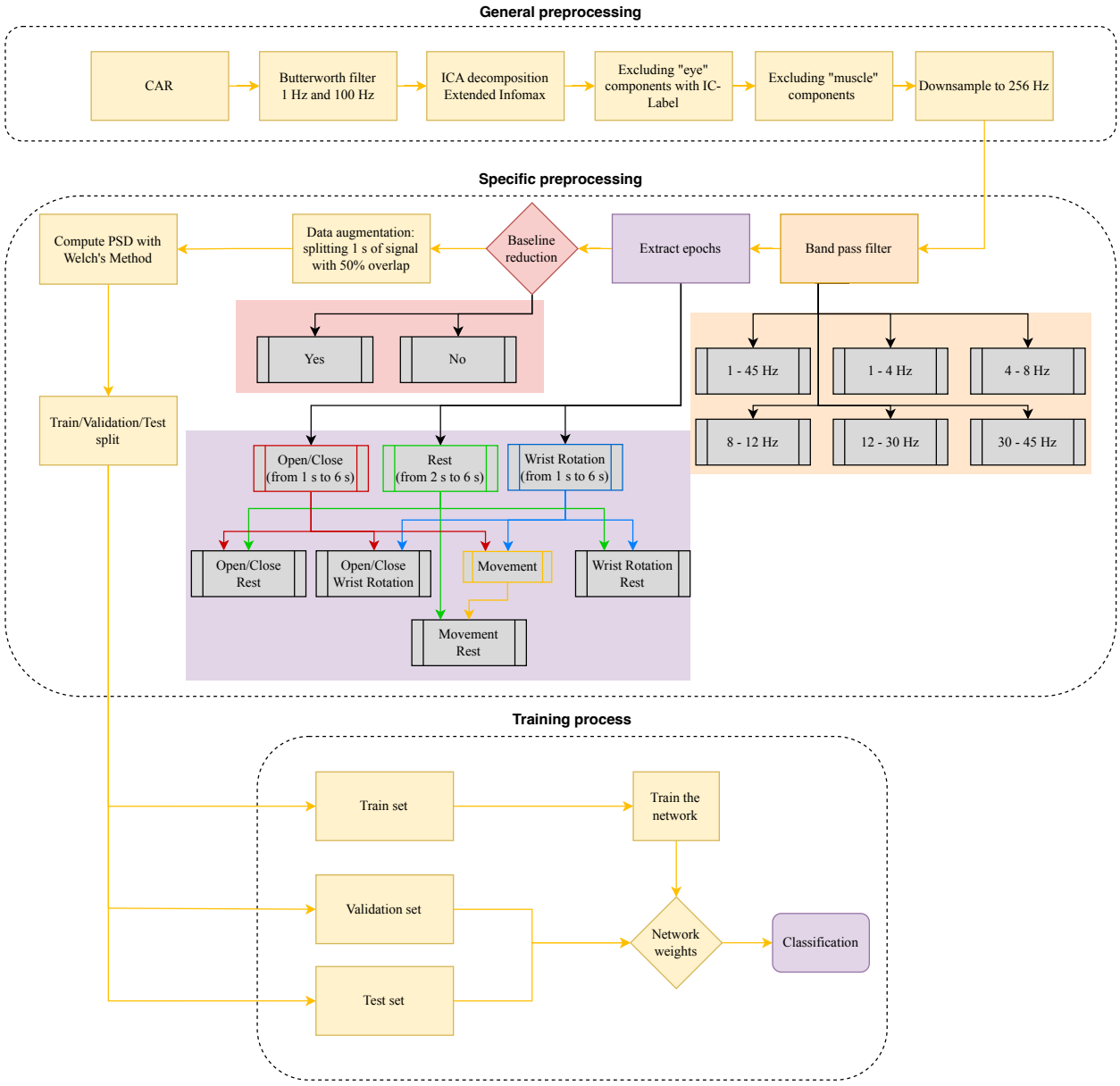
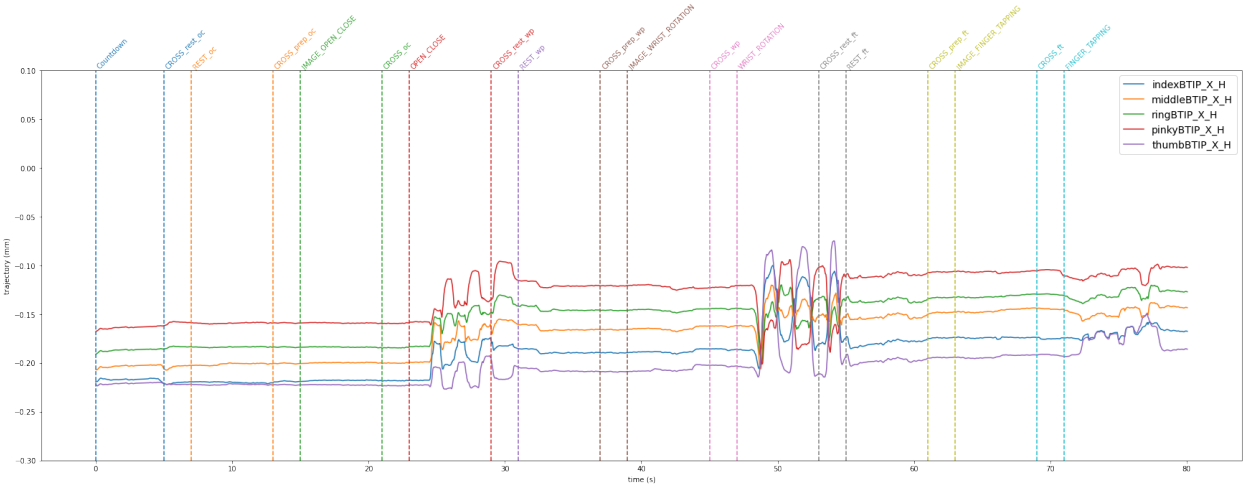
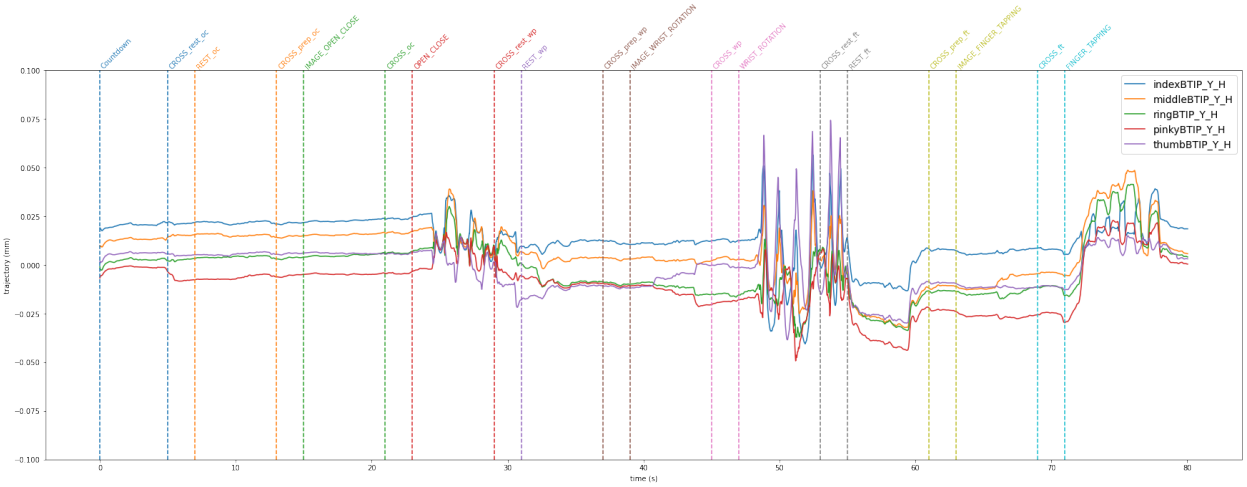


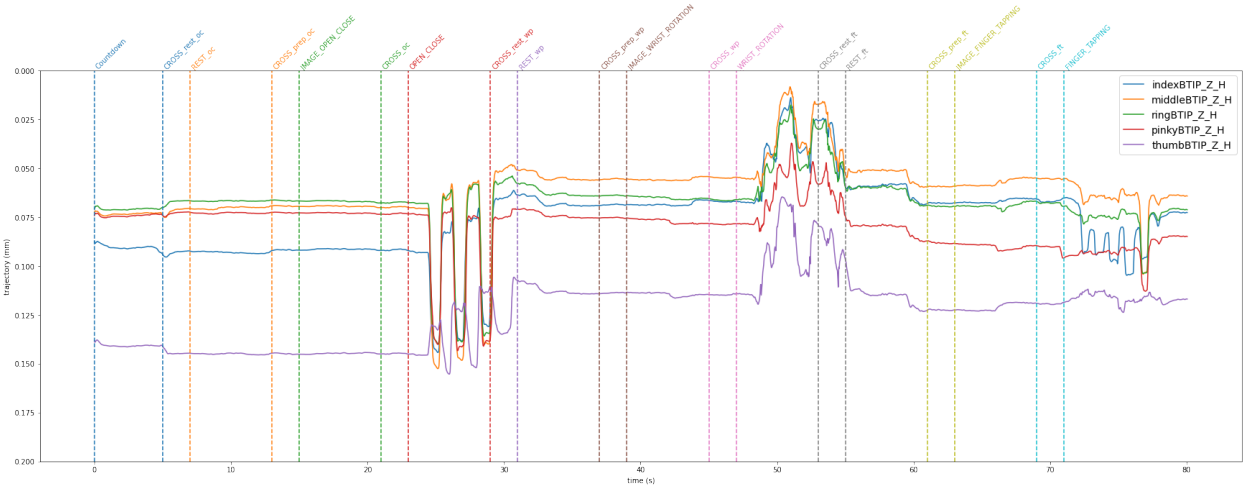
Figure 4.14: Preprocessing for sample generation and training. The colored boxes represent the parameters that change to explore different frequency bands (in orange) and the baseline reduction method (in pink), for each combination of movement/rest. The movement class (violet box) represents the merging of the open/close and wrist rotation. In the lower part of the scheme, the training process of the model is described.



(a)

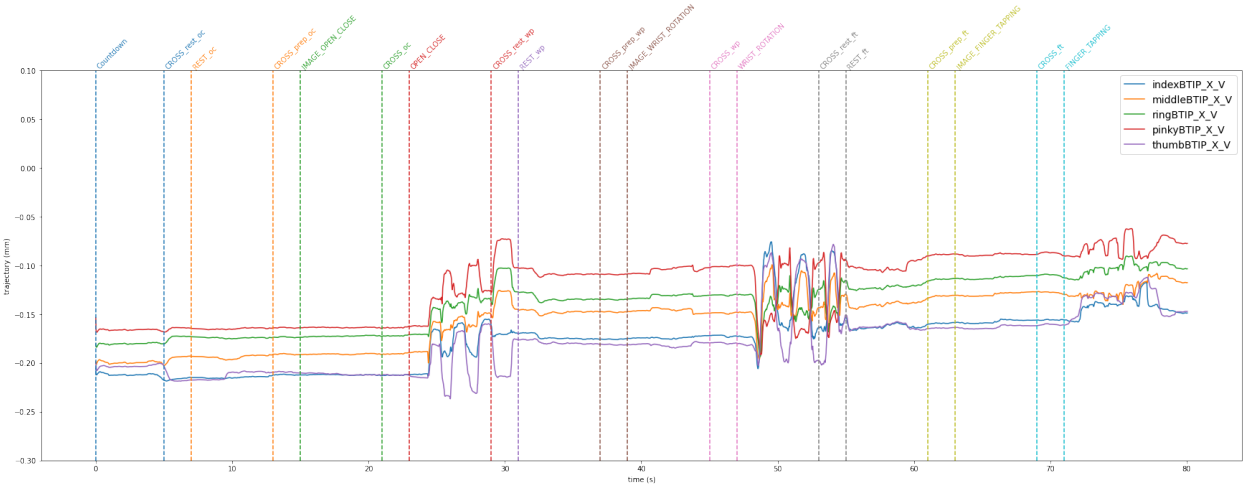


(b)

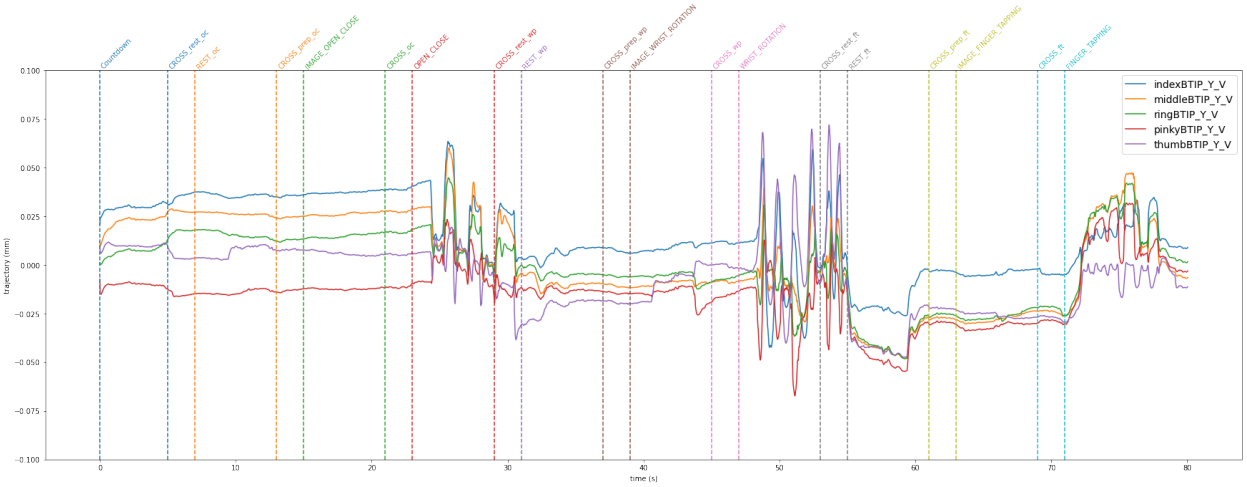


(c)

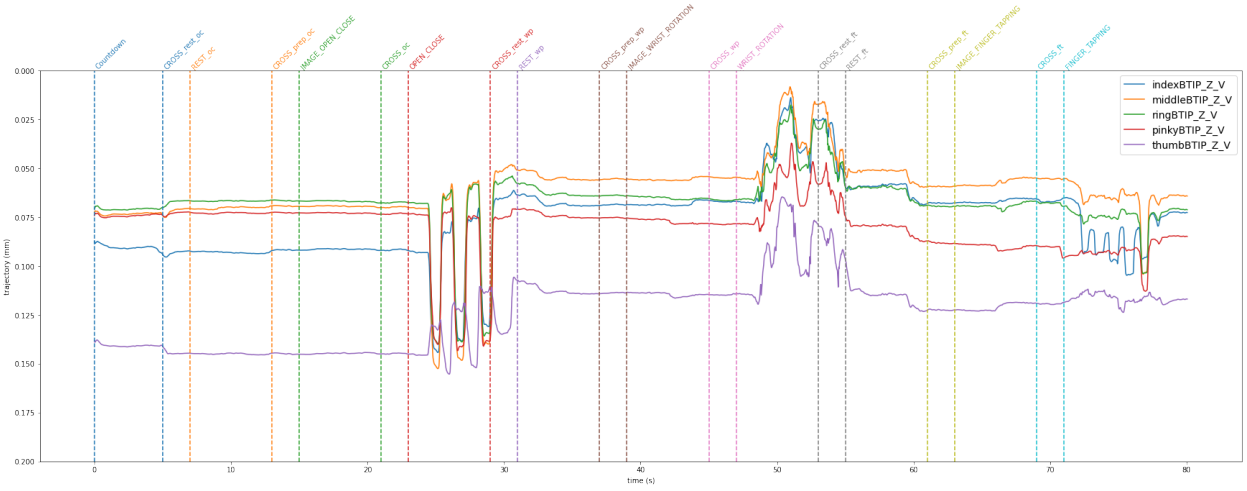
Figure 4.15: x (**a**), y (**b**), and z (**c**) components of the fingertip trajectory for the Horizontal LMC. All fingertips are reported in a single plot.



(a)

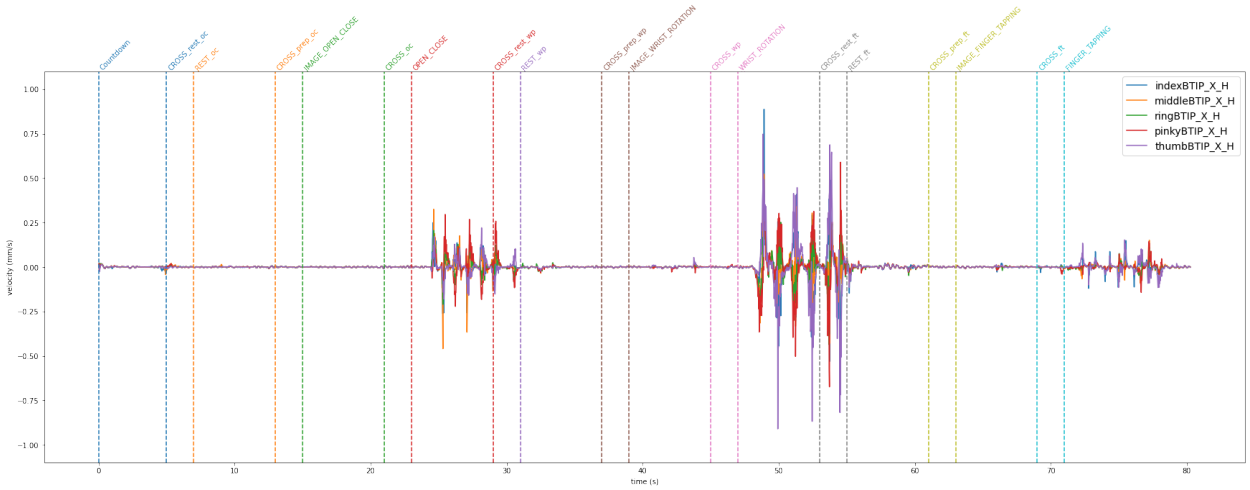


(b)

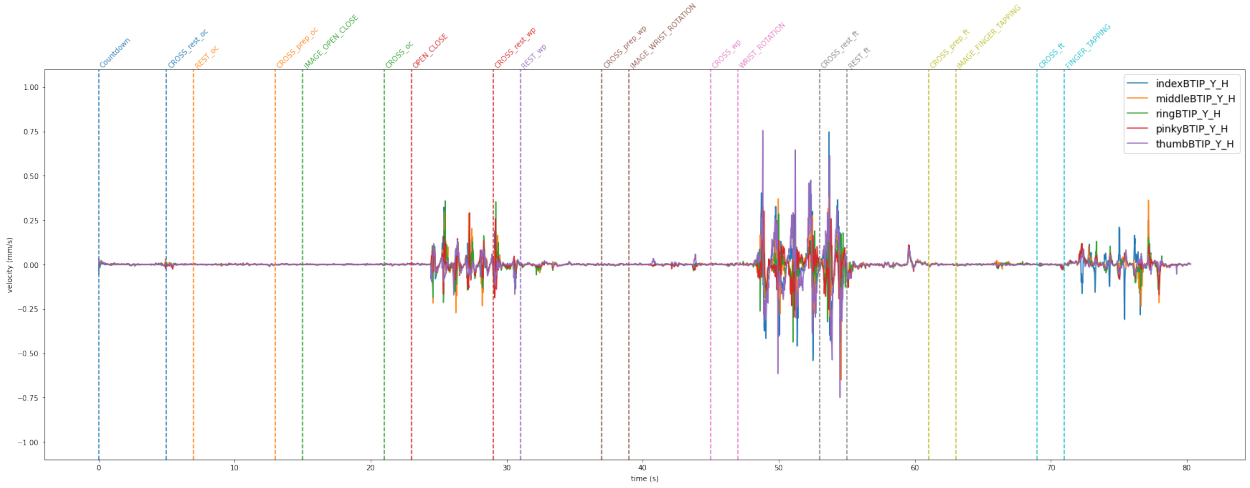


(c)

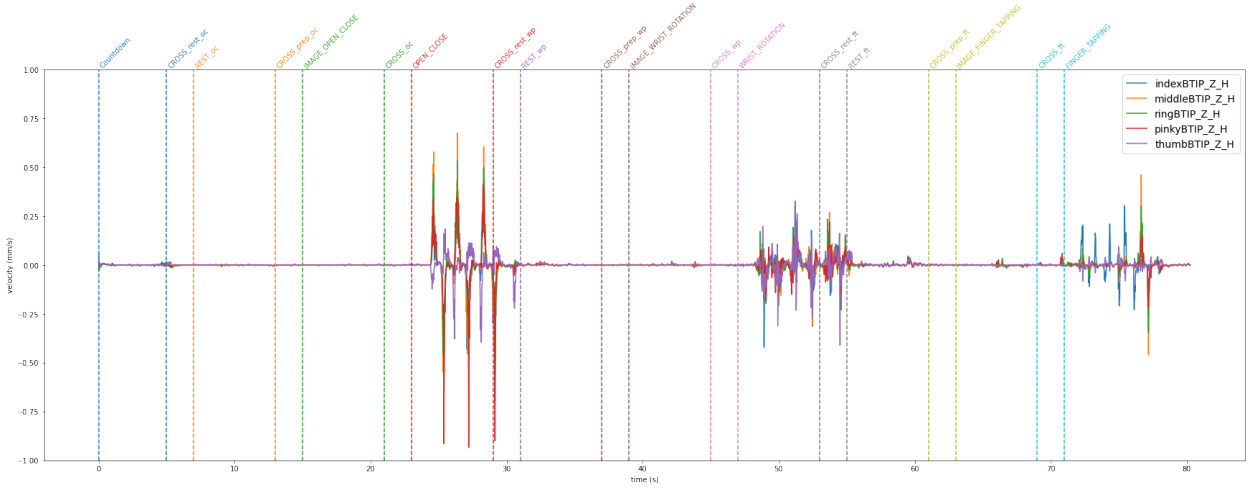
Figure 4.16: x (**a**), y (**b**), and z (**c**) components of the fingertip trajectory for the Vertical LMC. All fingertips are reported in a single plot.



(a)

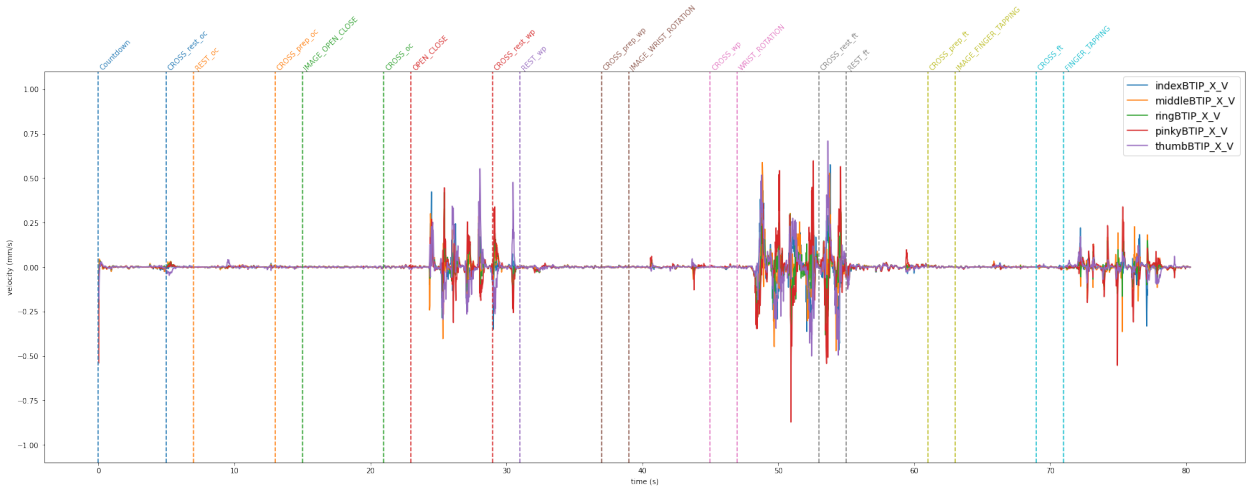


(b)

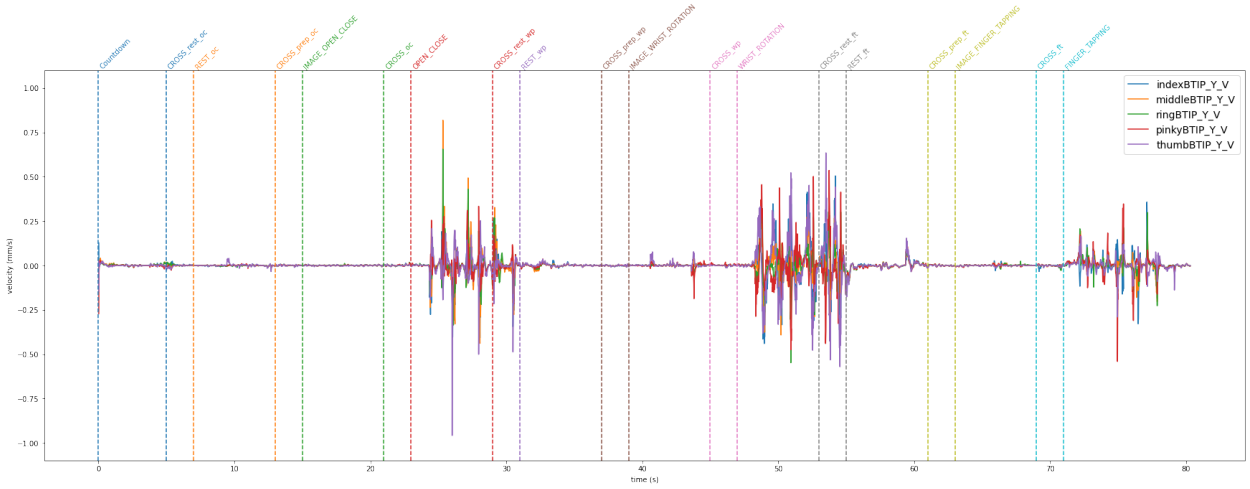


(c)

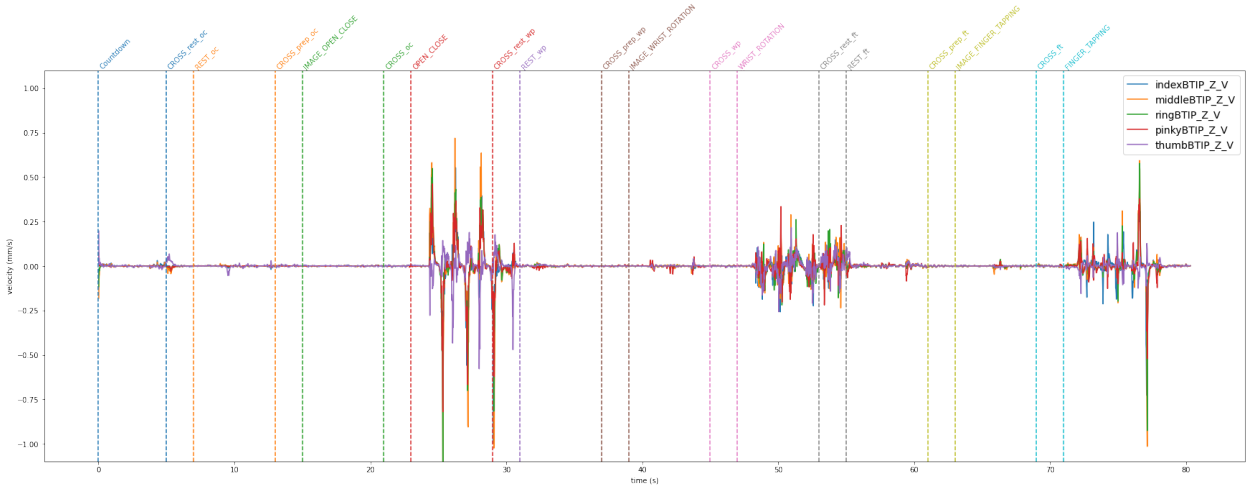
Figure 4.17: x (**a**), y (**b**), and z (**c**) components of the fingertip velocity for the Horizontal LMC. All fingertips are reported in a single plot.



(a)



(b)



(c)

Figure 4.18: x (**a**), y (**b**), and z (**c**) components of the fingertip velocity for the Vertical LMC. All fingertips are reported in a single plot.

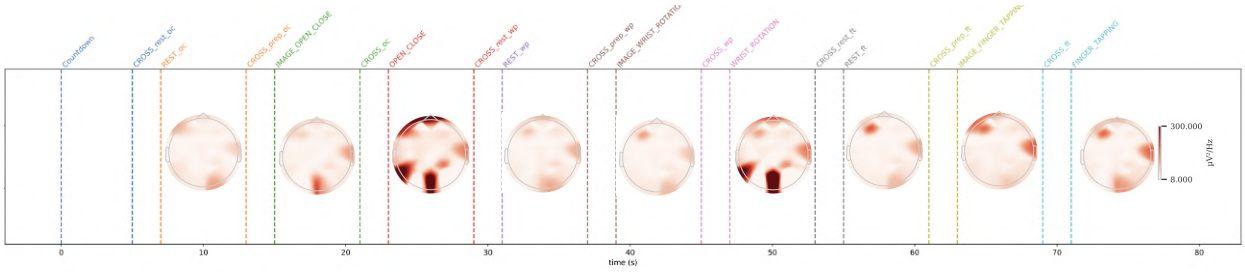


Figure 4.19: Spatial distribution of PSD for each task in a single subject. The timeline is the same as used in Figures 4.15–4.18.

Chapter 5

Sleep

The results of this chapter are published in:

D. Lozzi, A. Di Matteo, E. Mattei, A. Cipriani, P. Caianiello, F. Mignosi, G. Placidi, "**ASIS: a smart alarm clock based on deep learning for the safety of night workers**" in *IEEE MetroXRAINE 2024 - Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering*, 2024.
DOI: 10.1109/MetroXRAINE62247.2024.10796738

In this section the chapter focus on the analysis of the EEG data acquired during sleep. The main objective was to predict sleep states within 30 minutes of signal acquisition. The developed DL model, based on a novel architecture combining LSTM layers, allows the identification of the ideal time to wake an individual, maximising sleep duration, reducing sleep inertia upon awakening and improving cognitive performance.

Sleep inertia is a phenomenon characterized by drowsiness with motor and cognitive deficits susceptible to a state of intoxication [218]. It is characterized by a lower level of arousal [219] that occurs after awakening from sleep with a low level of cognitive performance [220] and its duration is around ~ 30 minutes, depending on chronotypes [221]. The arousal is defined as the level of physiological activation that is regulated by ARAS [4, 33, 222–224]. Moreover, the awake during SWS, where the SWA is high, which is either stage 3 or 4 of RK criteria or Stage N3 in AASM criteria, leads to a high level of sleep inertia. In contrast, awakening during stages 1, 2, and REM [219] (N1, N2 or REM in AASM criteria), leads to less level of sleep inertia [225]. Having less sleep inertia upon awakening makes a person more alert and responsive to external stimuli, which is essential for certain occupations such as truck drivers, air traffic controllers, and security guards [226], and the use of a method to reduce sleep inertia could improve the safety in many domains: the prevalence of die due to drowsiness are estimated around 21 % among all fatal crashes among car drivers [227]. To address this issue, it is important to identify the optimal time to wake up the subject when they are in the REM, N1, or N2 sleep stages (considered as "good" stages to wake up) where the percentage of SWA is low, while avoiding the N3 stage (considered as a "bad" stage to wake up) where the percentage of SWA is high. Additionally, it is necessary to predict the last good sleep stage before reaching the maximum duration of sleep, as the "good" sleep stages occur multiple times during the sleep cycle, but it is advantageous to utilize the last "good" stage until the maximum time is reached to make the sleep as long as possible.

The sleep stage scoring is classically performed by visually inspecting the EEG time-series during the night in semi real-time or offline mode, following the AASM scoring manual [228]. This approach is very time-consuming and heavily dependent on the researcher. Moreover, it is not possible to obtain the same results with absolute precision when using the same data and the same technician, due to the qualitative method used. Many models tried to automatically classify the sleep stage [229, 230] with very high performance. In

Sleep Stage	Description
Sleep stage 1	Light Sleep
Sleep stage 2	Sleep
Sleep stage 3	Deep Sleep
Sleep stage 4	Deep Sleep
Sleep stage R	REM (Rapid Eye Movement) Sleep
Sleep stage W	Awake
Movement time	Movement
Sleep stage ?	Unknown Stage

Table 5.1: Sleep Stages and Descriptions

[231–233] tried to classify the sleep stage according to EEG signals and using the PhysioNet Polysomnography dataset [234, 235]. In [236] they use the EDF-Sleep Expanded dataset to reach 82.0-76.9% of F1 score using CNN and bidirectional-LSTM to classify the sleep stages. An interesting review [237] analyzed 10 recents works on sleep stage classification based on EEG data using deep learning. In [238], [239] and [240], the authors used a CNN reaching 87%, 86% and 92% of accuracy respectively. In [241] and [242], were used an hybrid model composed by RNN and CNN with an overall accuracy on 80% and 91.3%. In [243], authors used another type of hybrid model composed by a MultiLayer Perceptron (MLP) and RNN with an accuracy of 83.6%. In [244], to help clinican on the sleep disease diagnosis, authors reached 90% of accuracy and in [245] authors used autoencoders to classify sleep stages with 89% of accuracy. In [233] the authors tested many models reaching an accuracy over 89% take into account also the unbalacned problem of sleep stages in the dataset used [234, 235]. In [246] the authors using an RNN model reached an overall accuracy of 78.90 $\pm$ 2.29%. W. Chen et al. also used an RNN model and reached an accuracy of 79.5%. In [247] and [248], the authors used SVM models to classify the sleep stages with precision 91.1% and 78.9%, respectively. The gold standard for classification is based on AASM [27], but many articles and many data sets are labeled according to the RK manual [25]. To our best knowledge, only two works have employed prior EEG data to forecast the future sleep stages. Choi and Sung [249] aimed to improve the accuracy of the current sleep stage classification by predicting the subsequent sleep stage. In [250], they used the data acquired with an actigraph, an instrument with a gyroscope that measures the movement of the subject, to predict the best moment to wake up. They used 138 sleep tracks acquired in 92 people to achieve 66-70% of accuracy classifying four sleep stages (deep, light, wake, REM) and 71-77% of accuracy in two classes (deep / REM vs. light / wake stages). To our best knowledge, no one used only two EEG signals to predict the best time to wake up the person to minimize sleep inertia condition, maximize sleep period, and optimize performance at the wake up moment. This could be useful to reduce accidents at work by improving physical and cognitive performance [30, 251–253] waiting as little as possible [254].

In this chapter, it was conceptualized the *ASIS - Anti Sleep Inertia System*, in order to identify the best moment to wake someone up using two EEG signals acquired with four bipolar referenced channels. This is done using a system composed of two machine learning models, (1) a *classifier* and (2) *predictor*. The predictor predicts all the best moments to wake up and the classifier checks the correctness of the prediction when the expected time is reached. The reason to use two models instead of one is that the classifier in general has higher level of accuracy, precision, and recall with respect to the prediction based on past data. The presented architecture is composed of two machine learning models that are able to predict the best moment, across a time range from a minimum of 30 seconds to max of 30 minutes (i.e. the lag between the last epoch used for prediction and the label predicted), to activate the alarm clock to have a less harsh awakening. Since there are many great models [237] already published that are able to classify sleep stage based on 30 seconds of signals using only four-channel bipolar referenced [233], in this chapter is presented only the (2) *predictor* part, and here the ASIS system is conceptualized using the one of the best classifier already published [233] and included in ASIS system. The predictor here presented uses the

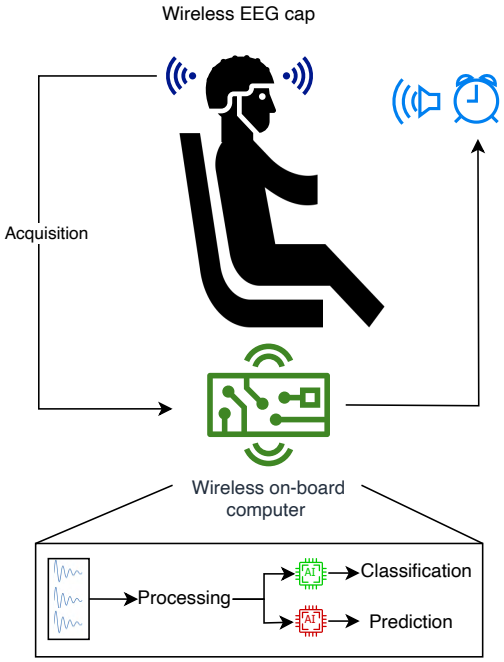


Figure 5.1: Use case of ASIS.

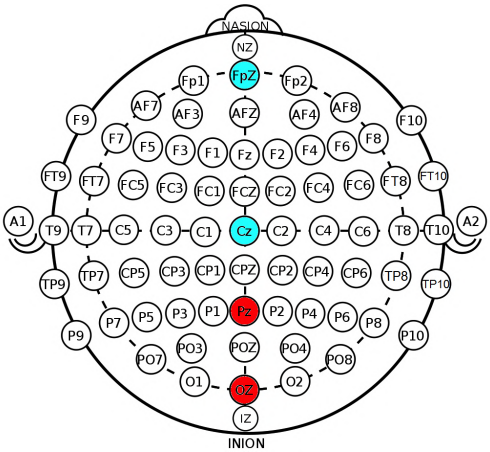


Figure 5.2: The 10-20 system electrode positions. The four channels with the same color record two signals based on two pairs of electrodes in a bipolar configuration.

AASM	% of the sleep AASM	RK	% of the sleep RK	Sleep Stage Characteristics	AASM stages description
N1	2 - 5%	Stage 1	2 - 5%	closed eyes, alpha rhythm	-
N2	45 - 55%	Stage 2	45 - 55%	spindles, k-complex	-
N3	7.5 - 15.5%	Stage 3	3 - 8%	20% slow wave activity, high voltage	slow-wave sleep (SWS), high
		Stage 4	10 - 15%	50% slow wave activity, high voltage	amplitude, low frequency
REM	25%	REM	25%	desynchronization, low voltage	-
Awake	0%	Awake	0%	desynchronization, low voltage	-

Table 5.2: Comparison between different classification systems [28].

Relative Band Power (RBP) from two EEG signals acquired with two couples of bipolar channels without any information about eye movement (EOG) or muscular activity (EMG). For the prediction phase, a LSTM model, a particular version of RNN is used to predict the future sleep stage based on an hour of signal with 30 minutes of gap from the input data to the label, as shown in Fig. 5.6 and in Alg. 3. Once the time reaches the first epoch predicted, the system checks if the label classified by the classifier matches what the predictor predicted. Within a preset time range, the alarm clock plays the sound of the last good epoch predicted and verified by the classifier. If the predictor does not forecast any good epoch or the classifier results do not match the label forecasted by the predictor, the alarm clock plays the sound at the upper limit of the time range.

5.1 Materials and Methods

5.1.1 The Sleep-EDF Database expanded

In this chapter, the preprocessed Sleep-EDF Database expanded was used [234, 235]. Data from 74 subjects sampled at 100 Hz were analyzed using MNE Library for EEG analysis [255]. Each subject contributed data from two nights of sleep recording, totalling 148 recorded nights of sleep. A bandpass filter (1-45 Hz) was applied to eliminate noise and DC current interference (60 Hz for the United States and 50 Hz for Europe). Each data contains the labelling performed manually by the original authors of the data set following the RK criteria. The datasets were manually labeled by the original authors using the RK criteria [25], and it was converted the original classification into one more close to AASM [27]. The EEG instrument used in the dataset was composed of four EEG channels with bipolar references (shown in Fig. 5.2): Fpz-Cz and Pz-Oz and other additional physiological channels (EOG channel, Resp oro-nasal, EMG submental, Temp rectal, Event marker).

5.1.2 Feature extraction

The ASIS utilizes EEG data from four channels that record two signals based on two pairs of electrodes mounted in a bipolar configuration (Fpz-Cz, Pz-Oz), excluding others. The length of each epoch for classification was set at 30 seconds, based on previous research and manuals [25, 27]. The analyzed signal was cropped from the initial drowsiness moment to the wake-up moment, excluding the remaining data. Subsequently, a list was created that includes the label, start, and end of each epoch in terms of the signal's timepoint index. The RBP was then computed for each segment and channel in the following six frequency bands: delta (1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), sigma (12-15 Hz), and gamma (30-45 Hz). RBP represents the average power within a specific frequency interval divided by the total power across all frequency ranges [53]. The average power was calculated by integrating the RBP using Welch's Method [99]. The use of RBP ensures that this information remains independent of the subjective value of the PSD for each subject [256]. However, it should be noted that the PSD during wakefulness may vary with age, as indicated in previous studies [257, 258].

In this chapter, it was use the RBP to be independent of subjective physiological difference in voltage [24]. The $BandPower$ is calculated by summing the $P_{xx}(f)$,where $P_{xx}(f)$ is the PSD of a signal $x[n]$, for each frequency of each band (Eq. 5.1), and the $TotalPower$ is computed summing all $P_{xx}(f)$ (Eq. 5.2).

$$BandPower_i = \sum_{b_i f=b_i f_{min}}^{b_i f_{max}} P_{xx}(f) \tag{5.1}$$

$$TotalPower = \sum_{f=f_{min}}^{f_{max}} P_{xx}(f) \tag{5.2}$$

Then, $BandPower$ and $TotalPower$ are used to compute $RelativeBandPower$, a shown in Eq. 5.3:

$$RelativeBandPower = BandPower/TotalPower \tag{5.3}$$

The result is the percentage of power for each band that is independent of the amplitude of the signal that cannot be stable between subjects. Furthermore, considering that $RelativeBandPower$ increases from 0 to 1, it is not necessary to scale the data [55] to train the ML models. Finally, a dataframe was constructed that includes information regarding the subject, recording, RBP for each electrode pair, and label.The pipeline for feature extraction is illustrated in Alg. 1 and the flowchart in Fig. 5.3.

In this study, the stages "Sleep Stage ?" and "Movement time" were excluded due to the small sample size. Ultimately, 153953 samples were extracted. Subsequently, Algorithm 2 was executed, as depicted in Fig. 5.4a and 5.4b. Considering that the labels are associated with the preceding 120 rows for each subject, and after excluding data related to "Sleep Stage ?" and "Movement time", the total dataset comprises 127313 samples. The distribution of samples is presented in Tab. 5.4 and 5.7. The EEG instrument captures the signals and sends them to the board depicted in Fig. 5.1, then the system functions as depicted in Fig. 5.6 and in Fig. 5.5: the acquired signals are cleaned with an artifact removal system [83, 85], and the RBP is calculated for every 30-second segment. After that, the classifier classifies each epoch based on data labelled as one of the sleep stages of the AASM classification. Once 120 epochs (equivalent to 1 hour of signals) have been acquired, the predictor begins its activity forecasting the sleep stage that will be acquired after 60 epochs (equivalent to 30 minutes) from the 120th epoch. Starting from the first 120 epochs acquired, the predictor continues the forecasting activity with each new epoch acquired, shifting the 120 epochs by one. With this shifting, the last 119 data already acquired are used, and the last segment of the signal is added to forecast the new label. Each sample is a matrix consisting of 120 rows representing time, and 12 columns representing the RBP in each band for each electrode. The label corresponds to the label of the 181th row of the data frame, which indicates the sleep stage after 60 epochs following the last row of the sample. The probability for each class, based on the number of samples, is presented in Tab. 5.3 shows the distribution based on the AASM classification: the small samples of classes 3 and 4 of RK classification, converge into the N3 category of the AASM classification and this is also done due to the spectrum similarity of classes 3 and 4 [24]. Therefore, it was opt to use a classification model that closely resembles the AASM classification. The training set to test set ratio was set at 0.33 and the validation was the 0.1 portion of the training set.

The Shapiro-Wilk test was applied to the RPB distribution across all channels and frequency bands, indicating that RPB does not follow a normal distribution. Following this, the Mann-Whitney test was performed across different classes and frequency bands, revealing statistically significant differences between the distribution of each class for each electrode in each frequency band. These findings supported the use of this feature in the predictor.

Algorithm 1 Feature Extraction Algorithm

for each raw data segment **do** ▷ Iterate over each raw data segment
    apply filter to raw data ▷ Apply a filter to preprocess the raw data
    for each epoch **do** ▷ Loop through each epoch in the filtered data
        for each channel **do** ▷ Loop through each data channel
            for each frequency band **do** ▷ Process each frequency band for the current channel
                P = extract power from filtered data ▷ Compute the power of the filtered data
                RBP = compute Relative Band Power (RBP) from P ▷ Calculate the Relative Band Power (RBP)
                append RBP to results dataframe ▷ Store RBP in the results dataframe
            end for
        end for
    end for
end for

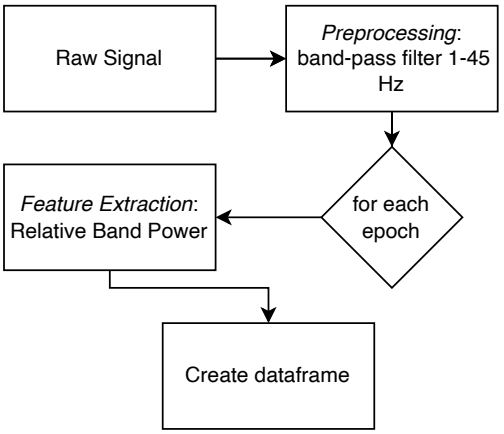
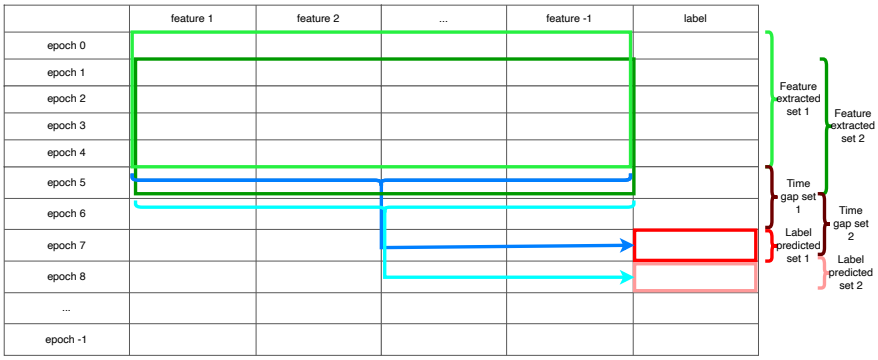


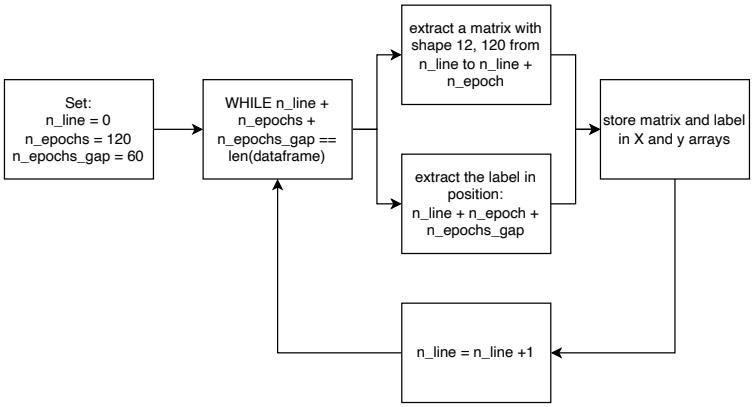
Figure 5.3: Feature extraction flow-chart used in this chapter.

Algorithm 2 LSTM Dataset Generator Algorithm

n_line = 0 ▷ Initialize the line index
n_epochs = 120 ▷ Set the number of epochs in each sequence
n_epochs_gap = 60 ▷ Set the gap between the epochs sequence and label
for each segment in dataset[0 : len(dataset) - (n_epochs + n_epochs_gap)] **do** ▷ Loop through each segment with sufficient length for LSTM
    for each epoch in dataset[n_line : n_line + n_epochs] **do** ▷ Iterate over the sequence of epochs
        extract RBP values for current epoch ▷ Retrieve Relative Band Power (RBP) features for the current epoch
        assign label from dataset at position n_line + n_epochs + n_epochs_gap ▷ Assign the label for the LSTM target from the dataset
        increment n_line by 1 ▷ Move to the next epoch in the dataset
    end for
end for



(a) LSTM dataset generator. A matrix of data are extracted, and the label associated with data are misaligned with the input data. The label of each input matrix data (i.e. the sample), correspond to the label of the 181 th row of the dataframe, starting from the 0th oh each input matrix data.



(b) LSTM flowchart dataset generator.

Figure 5.4: Dataset generator.

Algorithm 3 System overview

```
n_line = 0                                          ▷ Initialize the line index
n_epochs = 120                                      ▷ Set the total number of epochs to acquire
n_epochs_gap = 60                                   ▷ Set the gap between epochs for prediction
for each epoch do                                   ▷ Loop through each epoch
    append feature array to dataset                 ▷ Add feature array to the dataset
    classify epoch based on feature array           ▷ Classify the epoch based on the feature array
    while n_epochs_acquired == n_epochs do          ▷ If the required number of epochs has been acquired
        predict stage at position (n_line + n_epochs + n_epochs_gap)   ▷ Predict the stage at the specified position
        increment n_line by 1                       ▷ Increase the line index by 1
    end while
end for
```

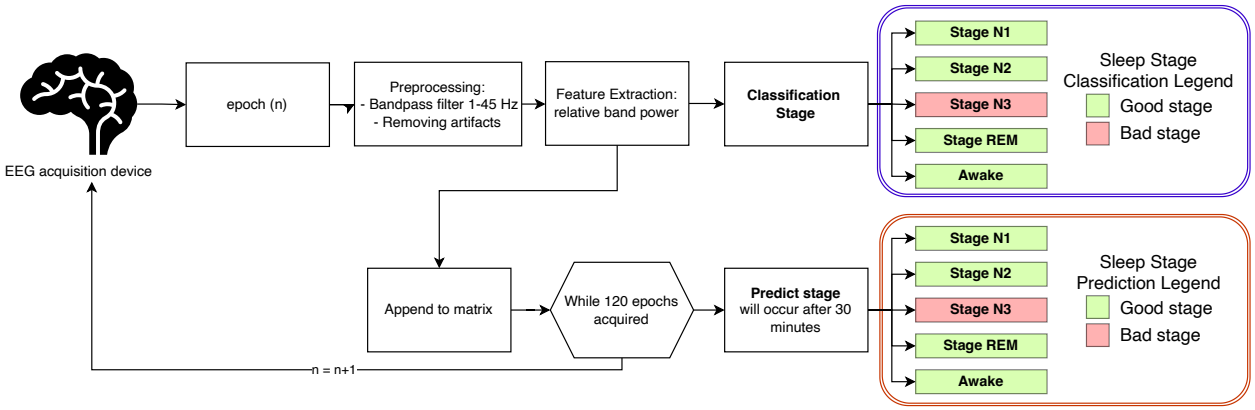


Figure 5.5: System classification and prediction. In this scheme, an artifact rejection algorithm is included because the ASIS is conceptualized for a real-time application [83, 85].

Sleep stage	**Probability**	**N. of samples**	**% sleep stage**
Sleep stage N1	0.100	15362	14.13 %
Sleep stage N2	0.395	60761	55.90 %
Sleep stage N3	0.052	7963	7.33 %
Sleep stage R	0.154	23651	21.80 %
Sleep stage W	0.294	45255	-
Movement time	3×10^{-5}	5	0.01 %
Sleep stage ?	0.006	956	0.88 %

Table 5.3: Probability of occurrence of each class in AASM classification.

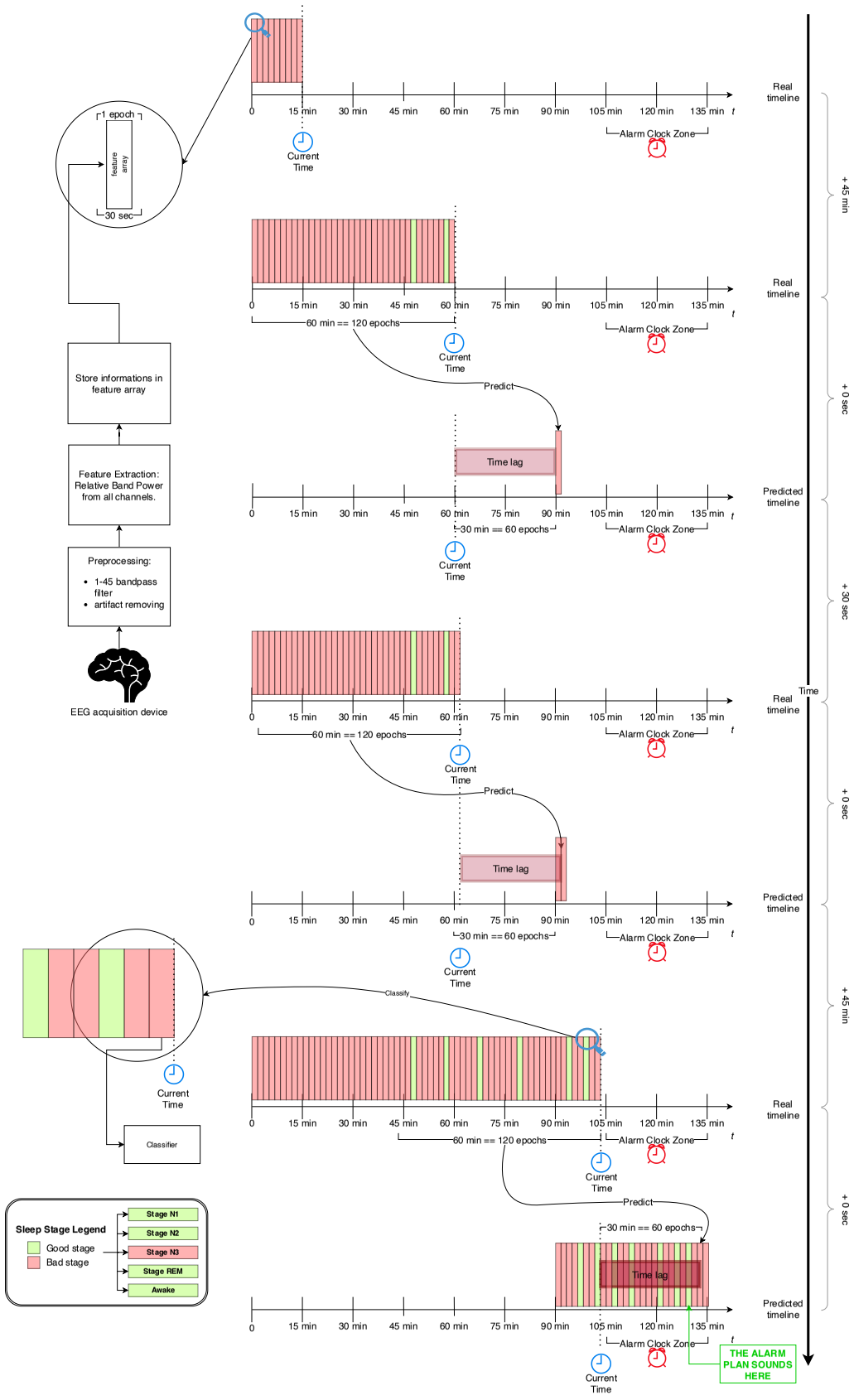


Figure 5.6: Description of the ASIS system. The prediction timeline stats its label forecasting and upon reaching the conclusion of the Alarm Clock Zone, the system rouses the subject at the last phase of light sleep to extend sleep duration and ensure an optimal wake-up. This suggests that the system delays waking until the most advantageous stage within the Alarm Clock Zone is achieved, if multiple appropriate stages exist. When the designated wake-up time arrives, the classifier assesses whether the current stage is appropriate for awakening. Should no appropriate stages be available, the alarm is set off at the maximum boundary of the Alarm Clock Zone. Note: the array sizes depicted do not match 30-second intervals on the timeline; they are merely illustrative of the procedure.

Model	LSTM
Total sample set	127313
After Sleep stage exclusion	126352
Balanced	26490
Training	17748
Test	8742
Validation	10 % of training

Table 5.4: Sample number of the training set for LSTM.

Performance on prediction with LSTM on AASM labels

	Precision	**Recall**	**F1-score**	**Support**
Sleep stage N1	0.60	0.70	0.65	1733
Sleep stage N2	0.65	0.53	0.58	1805
Sleep stage N3	0.84	0.91	0.87	1682
Sleep stage R	0.74	0.76	0.75	1758
Sleep stage W	0.84	0.79	0.81	1764
Accuracy			0.74	8742
Macro avg	0.74	0.74	0.73	8742
Weighted avg	0.74	0.74	0.73	8742

Table 5.5: Performance on Prediction with LSTM on AASM labels.

5.1.3 The predictor

The prediction of the label is carried out using preprocessed data. To ensure that the back-propagation does not encounter any issues, the classes in the dataset are balanced. This is done by resampling the other classes based on the class with the fewest number of samples. The Tab. 5.7 shows the results of this balancing process. Random samples are taken from the original set during the pruning process. In order to avoid creating an imbalance problem due to the low number of samples in the "bad" sleep stages, the prediction is not performed by merging all classes into only two classes. Merging the N1+N2+R+W stages and leaving N3 alone would significantly reduce the number of samples in the balancing stage. Furthermore, the significant differences in the RBP between each sleep stage make the merging approach less effective for the prediction. This is because the "good" stages do not have any spectrum similarity, but share the property of having less sleep inertia at awakening. To reach this goal, an LSTM composed of 6 layers was built. This type of RNN addresses the issue of gradient vanishing by incorporating three gates on each memory cell: input, forget, and output. These gates enable the model to capture the temporal dependencies in sequential data. The activation of the cell is regulated by these gates, which control the flow of information through the cell by selectively updating the hidden state and determining which information to discard or output [259]. The architecture is depicted in Fig. 5.7, which is designed to be lightweight. Subsequently, the model was trained using a balanced training set consisting of 17748 samples. Each sample comprised 120 rows representing *time* and 12 columns representing features, and each label is associated with the 181th epoch. The test set contained 8742 samples, with the same shape. The parameter for the validation set was set to 10 % of the training set. Taking into accout the Tab. 5.7, the results depicted in Fig. 5.8 shown an high accuracy over all classes, and the low percent of error shared by similar sleep stages are around 10 %. Form the Tab. 5.5, the model shows a good classification performance in each classes over each metrics, and also in the overall performance shown in Tab. 5.6. The model shows promising results in the prediction phase, reaching an overall accuracy of 0.73, compared to $\sim$ 0.20 chance level [145].

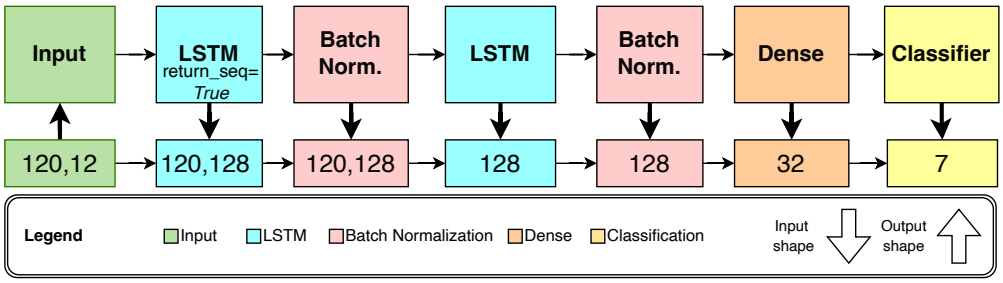


Figure 5.7: LSTM network architecture

Metrics for prediction using LSTM

Metrics	**Value**
Accuracy	0.7353
Balanced Accuracy	0.7378
F1 Score	0.7331
Recall Score	0.7353

Table 5.6: Metrics for Prediction with LSTM.

	?	N1	N2	N3	R	W
Before Exclusion "?"	961	5298	50110	21458	23651	37505
Chance	0.007	0.038	0.361	0.153	0.170	0.270
Not Balanced	-	5298	50110	21458	23651	37505
Chance	-	0.038	0.363	0.155	0.171	0.271
Balanced	-	5298	5298	5298	5298	5298
Chance	-	0.2	0.2	0.2	0.2	0.2

Table 5.7: Distribution of the stages in the dataset.

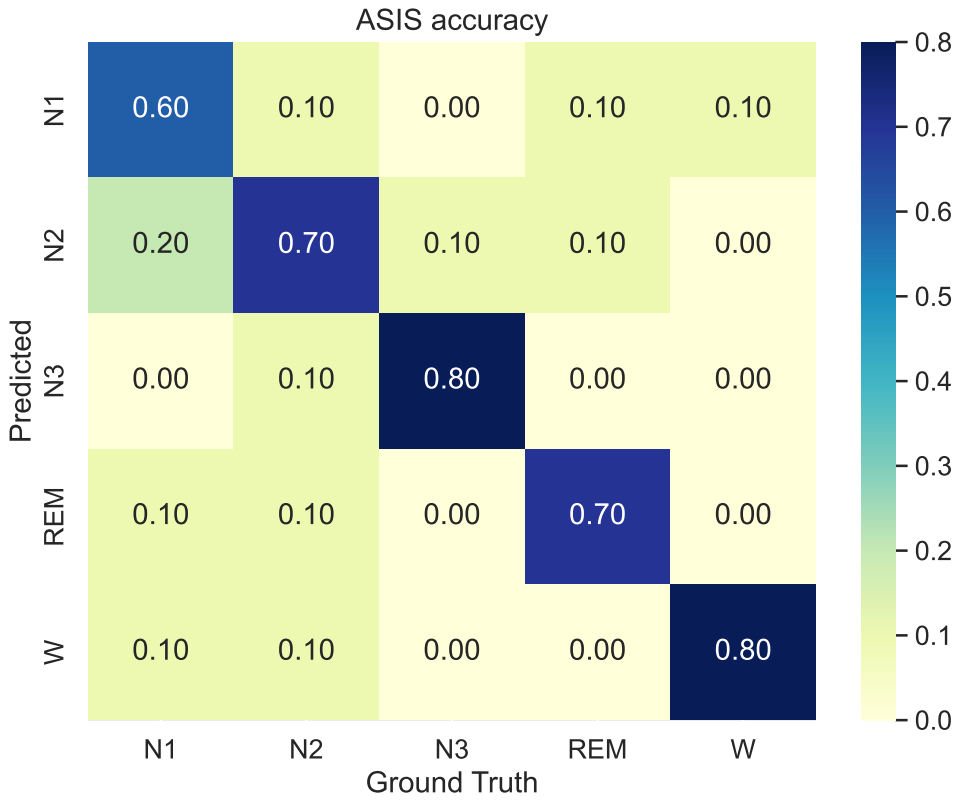


Figure 5.8: Confusion matrix for prediction stage.

Channels	*Fpz-Cz*	*Pz-Oz*
Features extracted	Fpz-Cz_delta	Pz-Oz_delta
	Fpz-Cz_theta	Pz-Oz_theta
	Fpz-Cz_alpha	Pz-Oz_alpha
	Fpz-Cz_beta	Pz-Oz_beta
	Fpz-Cz_sigma	Pz-Oz_sigma
	Fpz-Cz_gamma	Pz-Oz_gamma

Table 5.8: List of features used.

5.2 Descriptive Statistics

In the Appendix, the comparison between channels, frequency bands and sleep stages in function of their RBP in order to support the use of this features for sleep stage classification and prediction. The list of features is shown in Tab. 5.8. In the appendix, Tab. A 2 and A 4 display the average RBP for each category of the AASM classification criteria. The data is divided by sleep stage, channels, and frequency band in Tab. A 2, and by channel, sleep stage, and frequency band in Tab. A 4. In Tab. A 3 and A 5, the results of the Mann-Whitney test are presented. The test was conducted between each class, divided by channels and frequency band in Tab. A 3, and by channel, sleep stage, and frequency band in Tab. A 5. The p-value is included in each cell, indicating that the differences in the distribution of RBP values between each class are statistically significant. The Mann-Whitney test was chosen after performing the normality Shapiro-Wilk test, the results of which can be found in Tab. A 1. Moreover, in Tab. A 6, A 7, A 8, and A 9 illustrate the distribution of RBP values with and without normalization, providing a better visualization of the data in our dataset. Finally, Tab. A 10 presents the spectrum for each channel and each class of the AASM classification, clearly demonstrating the distinct behavior of each sleep stage.

5.3 Results

The objective of this study is to introduce the idea of the models employed in the ASIS. This architecture aims to improve the efficiency of alarm activation within a specific timeframe, while also maximizing sleep duration and minimizing sleep inertia. This is achieved by predicting the optimal time to wake up and evaluating the accuracy of the prediction using a classifier during the predicted epoch. These two models operate simultaneously: the classifier starts its operation after 30 seconds of detecting drowsiness, while the predictor starts its operation after 120 epochs (equivalent to 60 minutes) since it requires 120 batches of data to forecast the sleep stage after 30 minutes of the last epoch. The LSTM model shows promising results in the prediction phase, reaching an overall accuracy of 0.73, despite its lightweight architecture as shown in previous work [2]. To enhance the performance of the classification and prediction neural network, future research should explore different features and employ Leave-One-Subject-Out cross-validation to account for subjective differences. In real-world scenarios, it is essential to use artifact rejection systems [83] to ensure the acquisition of clean data and analyzing using system to minimize the latency for between the server and the user [260]. Moreover, considering the dynamic nature of the sleep cycle throughout the night, future work may include information about the specific time point in the night when the subject is being monitored and the approach of graph analysis [4] or other processing techniques for nonstatioanry signals [93], and incorporating additional robotic [200] or sensor systems [5] to enhance the human-computer interaction experience. Finally, could be interest to assess the sleepiness before and after using ASIS system [261].

Chapter 6

Emotions

The results of this chapter are published in:

D. Lozzi, E. Mattei, R. Ciuffini, A. Di Matteo, A. Marrelli, R. Ornello, M. Polsinelli, C. Rosignoli, S. Sacco, G. Placidi, "**The challenge of driving BCI with emotional signals collected by EEG**", *9th Graz Brain-Computer Interface Conference 2024*, pp. 366–371, 2024. DOI: 10.3217/978-3-99161-014-4-064

D. Lozzi, F. Mignosi, G. Placidi, M. Polsinelli, "**Graph model of phase lag index for connectivity analysis in EEG of emotions**", *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 348–353, 2023. DOI: 10.1109/CBMS58004.2023.00242

D. Lozzi, F. Mignosi, M. Spezialetti, G. Placidi, M. Polsinelli, "**A 4D LSTM network for emotion recognition from the cross-correlation of the power spectral density of EEG signals**", *2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pp. 652–657, 2022. DOI: 10.1109/WI-IAT55865.2022.00103

D. Lozzi, E. Mattei, A. Di Matteo, C. Rosignoli, F. Mignosi, G. Placidi, "**REMEMO: REMembrance of EMOtions**", *in preparation*

The section 6.4 "**Neural correlations of visual emotional processing**" are derived from the Erasmus/Visiting Ph.D. experience at the Institute of Neural Engineering (Technische Universität Graz) under the supervision of Prof. Wriessnegger[1].

Part of the code used for the Section 6.3 is reported in the Appendix D.

This chapter presents the latest experiments conducted based on preliminary studies with EEG signals acquired in movement and sleep. In particular, the literature and different datasets of public emotional EEG were first examined. Subsequently, innovative architectures were explored for the classification of the emotional state of relax, comparable to the "rest" state in movement. Consequently, innovative methods for the analysis of functional connectivity were studied and tested. Finally, based on the results obtained, two experimental protocols were developed for the acquisition of EEG data in two different elicitation contexts: the first active, where subjects were stimulated with images with high emotional content, and the second passive, where subjects were asked to recall emotionally activating events. The analysis were carried out with different AI methods that highlighted the intrinsic difficulties and biases present in the acquisition of data and in the classification of EEG signals with emotional content.
Emotion recognition using EEG is considered a promising tool for understanding human emotional responses,

[1] https://www.tugraz.at/institute/ine/research/team-wriessnegger

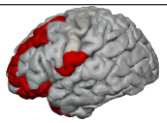
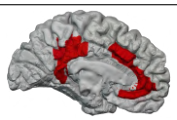
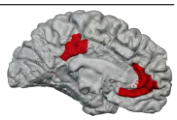
Emotion	Factor 1	Factor 2	Factor 3
			
Anger	-0.71	0.00	0.76
Happiness	1.68	0.54	-0.83
Disgust	-0.60	2.03	0.00
Sadness	-0.84	-0.59	-1.46
Fear	-0.69	0.22	-1.10

Table 6.1: Factor underline emotion related activation [273].

with applications in Affective Computing and BCI systems. However, this methodology faces challenges and limitations that require careful analysis. Since the amygdala is a subcortical structure, the in vivo study of emotions requires techniques capable of acquiring direct (e.g. EEG) or indirect (e.g. fNIRS) data on its functional activity. In recent decades, numerous studies published in peer-reviewed journals have shown promising results in classifying EEG signals using AI algorithms ,ML, DL, and transformers), achieving metrics such as accuracy, sensitivity, and specificity exceeding 90 % in multiclass classifications. However, a significant criticism is the notable interindividual variability of EEG signals, influenced by factors such as fatigue, cognitive load, and physiological brain differences [262]. This variability limits the generalizability of the results in diverse populations, which complicates the development of robust and reliable models [263]. Furthermore, the intrinsic complexity of emotions, which vary according to context and individual, poses further challenges [264]. Another critical issue concerns the quality of EEG signals, often contaminated by noise and artifacts, making it difficult to isolate components specifically related to emotions [10]. Many studies focus on single-electrode analysis, overlooking the importance of functional connectivity among different brain areas, which could provide more comprehensive and meaningful information [266]. Furthermore, the lack of systematic strategies to address noise in data can lead to overestimated or inaccurate results. Although the use of advanced techniques such as deep neural networks has improved model performance, the complexity of EEG data requires even more sophisticated approaches to ensure reliable outcomes [267]. Another significant concern is the ecological validity of the results: many experiments use artificial paradigms, often far from everyday emotional experiences, limiting the applicability of models in real-world scenarios [268]. Therefore, it is crucial to develop systems that can operate in natural and dynamic contexts [269]. Despite these challenges, emotion recognition via EEG remains a highly compelling area of research. Addressing issues such as interindividual variability, data quality, ecological validity, and model generalizability is essential to improve the utility of these technologies. Based on biological and anatomical knowledge of CNS, EEG data do not allow for the direct identification of pure emotional states but reflect the activation of the cortical structures involved in emotional processing, as discussed in Chapter 2. This limitation arises from the nature of EEG, which primarily records the activity of the upper cortical layers, as explained in Chapter 3. Despite the source reconstruction techniques described in Section 3.3.2, the difficulty of isolating signals originating from small, interconnected subcortical structures, combined with the presence of artifacts, complicates the identification of valid emotional patterns at the levels between subjects and between subjects. This complexity requires careful attention to interpret and generalize the results. Following recent developments in the literature [137, 270, 271], the objective of this study is to classify emotions by distinguishing them from a *relax*. Identifying a neutral state is crucial for a BCI classifier to recognize when *not* to perform any action, as the user is waiting for a stimulus or in a *rest* phase, similar to what occurs in motor BCI systems [7, 9].

Emotion recognition is part of Affective Computing [274], but is also used in other domains [2, 4, 264, 275–278]. and recently the interest of the scientific community has grown, as shown by the graphs in Fig. 6.1, emotions being one of the main aspects of human activity and mutual interaction.

The identification of the area related to each emotion is well studied by [273]. In this section, the

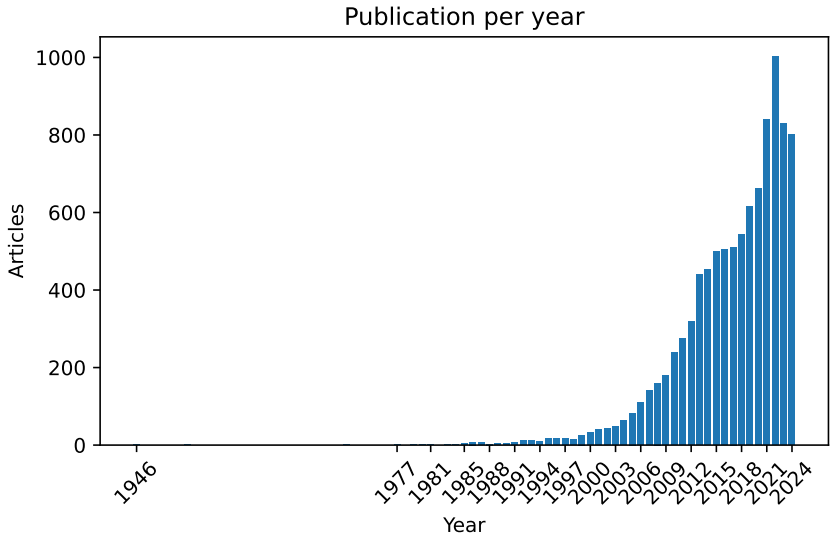


Figure 6.1: Trend of the number of publications by searching, in the PubMed database, the terms *"Emotion" AND "Recognition"* in *"article title" OR "abstract"* (searching date: January 1st, 2025).

authors used fMRI to record self-inducted emotions by actors, identified some neural signature of emotions, extracting the values of each factor associated with each emotions. The results are shown in Tab. 6.1. In recent years, attention has also been focused on the network of interconnected regions involved in emotional processing, even if neuropsychological data do not confirm the theory of a single emotional network, but of multiple networks that control multiple emotions [32]. In these networks, structures such as the thalamus, the somatosensory cortex, the somatosensory association cortex, the amygdala, the insula, and the medial prefrontal cortex have been clearly recognized [32]. Actually, besides the lack of an univocal definition of emotion, there is also no consensus in the scientific community on a general theory of emotions. In fact, the theory of emotions is divided into two main models: discrete and dimensional models [279, 280]. Discrete models classify emotions into categories without considering any axis to quantify the specific characteristics of each. Furthermore, the number of emotions depends on the theoretical framework. Dimensional models define some common continuous features and place each emotion in a point in the space based on the values of the considered features.

The study of the brain response to emotions and emotion recognition is useful in several applications, such as the diagnosis of mental disorders, helping the physician understand the emotions of some patients, the implementation of empathetic systems or driving a BCI, also for people with disabilities [281, 282], and EEG is probably the best technique to measure brain activity in real time and under daily conditions [281]. The use of physiological signals avoids subjective rating and is more reliable, but most of the measurement strategies listed above are too invasive and bulky to measure emotions in daily life conditions. Among them, EEG is one of the most indicated in assessing emotional states under almost normal conditions [281]. Generally, experimental works aim to classify many emotions and reduce the complexity of performing binary classification. However, an accurate study of emotions would require the measurement of emotions. Measurement of spontaneous emotions, under daily living conditions, is also impossible for EEG and, usually, external stimulation sources (videos, sounds, images) are used to elicit emotions in well-defined experimental sessions and environments. The resulting signals are stored in public datasets to allow open-source recovery, processing, and analysis, also in automatic ways. Most of the results produced are described in a dimensional space of emotions and few studies have been conducted on the categorical classification of emotions [283]. Several works aimed to discover specific patterns of EEG signals with automatic methods, also based on AI [182, 278, 283–285]. However, to understand the brain processes that underlie emotions and to improve the robustness of interpretation, the collected data sets should be huge and composed of labeled data. In

fact, emotion recognition is a particularly difficult domain because of the number of variables influencing emotions, including the elicitation mode, the experimental protocol, the subject's basal emotional state, and the instrumentation used.

6.1 The challenge of Emotion Recognition

This section describes the important challenges of driving a BCI through EEG signals of emotions. In particular, the complex emotional processing activated by the human brain and the need to generate elicitation protocols to synchronize the acquisition of EEG signals from emotions are presented. In addition, the limitations of EEG in dealing with signals from emotions are also discussed. Then, specific neuropsychological issues related to the use of protocols to elicit emotions are described. Due to the huge difficulty in managing the uncertainty deriving from the above issues, the surprising results obtained by recently proposed automatic strategies for emotion classification and recognition, also raising doubts about the correctness of the results, are reported and discussed. Finally, suggestions are presented on some procedures for the reduction of uncertainty and for the complete development of emotional BCI based on EEG.

BCIs, mainly those used as an alternative communication tool for disabled people, are based on event-related signals induced by external stimuli and synchronized with them (an example is theP300 [286]). Another consistent part is based on sensory–motor rhythm amplitudes [287]. For patients with impaired vision, or suffering from seizures attacks caused by too fast visual stimuli such as those used inP300, or that have never experienced the control of the motor part of their body, or whose signals produced by sensory–motor rhythms, mostly at the alpha band, can be easily confused with those due to artifacts caused by involuntary and frequent movements, other paradigms, such as auditory [288] and tactile [289], have been explored. However, cases in which also these paradigms have little or no effect are frequent [288]. Moreover, there are some situations in which some well-known techniques to build a BCI are not possible, like severe brain injuries that affect, for example,P300 elicitation [290]. For this reason, new ways have been explored and one of the most promising is that BCIs are based on the voluntary brain activity produced by emotions, being emotions mostly related to the deeper parts of the brain, mostly unaffected by disabilities [2, 291–295]. Emotions were first explored in the field of affective computing where some fascinating studies are dedicated to making the computer more empathic to the user and involved in the measurement of the user's emotions and representing them into HCI systems [296]. They aim to find the activation of specific brain regions in response to specific emotions but, while some regions are more active than others when experiencing specific emotions, no specific region is activated by a single emotion [273, 297, 298]. The brain regions most responsible for emotions are the amygdala, insula, anterior cingulate cortex, and orbitofrontal cortex. Through fMRI, it has been found that there exist specific patterns of brain activity, i.e. groups of brain districts, related to specific emotions and that these patterns are common across individuals [273]. A scheme of the brain regions mostly involved in the processing of emotions is indicated in red in Fig. 6.2. Despite these advances, it remains very difficult to recognize emotions, especially across individuals, because their patterns are very similar and can confuse each other and because they are also subjective (i.e. different individuals can have different ways of dealing with emotions). Moreover, complex multi variable pattern analysis techniques have to be used to identify distributed patterns associated with specific emotions, especially from EEG. Indeed, in EEG the sensitivity of the electrodes is higher for external neurons and decreases for deeper neurons (yellow part in Fig.6.2). This means that not all neurons equally contribute to the EEG signal, with an EEG predominately reflecting the activity of cortical neurons near the electrodes on the scalp. Deep structures within the brain, mainly involved in the emotional process, are further away from the electrodes and will not contribute directly to an EEG. Furthermore, EEG signals have an intrinsic nonlinear and nonstationary nature. The basic source of the nonstationarity in EEG signal is a reflection of switching of the inherent quasi-stable states of neural assemblies during brain functioning and is not due to the casual influences of the external stimuli on the brain mechanisms. EEG signal recorded from a scalp

electrode is influenced by different deeper sources, each 'transmitting' with different and variable intensity, thus making the main source of the registered signal from one brain structure to another. Nonstationarity also arises because of different time scales of the dynamic processes of brain activity. Finally, the signals are affected by noise and artifacts which, in some cases, are difficult to reduce effectively. Despite that, recently proposed nonlinear automatic strategies, learning by data collected during task-related protocols execution, promise to recognize emotions with very high accuracy, starting from EEG signals collected by eliciting protocols. In what follows, first the complex emotional processing activated by our brain, the disagreement in creating a commonly accepted model for representing emotions, and the necessity of generating elicitation protocols to synchronize the acquisition of EEG signals from emotions is discussed. Then the specific neuropsychological issues related to the use of protocols for eliciting emotions is described. Moreover, it is emphasize the surprising results obtained by recently proposed automatic strategies for emotion classification and recognition from EEG signals collected by these elicitation protocols, also raising doubts about the correctness of the results in light of the great uncertainty of the data and protocols. Finally, some cues for the future of EEG-based emotional BCIs is proposed.

6.1.1 Brain emotional processing and emotional models

The neural locations involved in the genesis and processing of emotions are multiple, including Autonomic Nervous System (ANS), the hypothalamus, the ascending reticular system, the limbic system, some lobes of the cortex, and the amygdala [32]. In recent years, attention has also been focused on the network of interconnected regions involved in emotional processing, even if neuropsychological data do not confirm the theory of a single emotional network, but of multiple networks that control multiple emotions [32]. In these networks, structures such as the thalamus, the somatosensory cortex, the somatosensory association cortex, the amygdala, the insula, and the medial prefrontal cortex have been clearly recognized [32]. Although there is no single shared definition of emotion, many researchers define an emotion as a feeling related to an event that occurs during a subjective experience characterized by a complex brain function that includes information acquisition, manipulation, storage, and recall [299]. Ochsner and Gross [300] assumed that emotions are consequences of external stimuli or internal mental representations with valence and well-defined characteristics. Furthermore, some researchers [32] hypothesize that emotions have three components: physiological reaction, behavioral response, and feeling, i.e. a subjective response to emotions. Emotion recognition aims to detect the affective state of a subject directly by brain activity, and recent work focused the attention on the correlation between brain oscillations and emotions [182, 301]. In fact, besides the lack of a univocal definition of emotion, there is also no consensus in the scientific community on a general theory of emotions. In fact, the theory of emotions is divided into two main models: discrete and dimensional models [279]. Discrete models classify emotions without considering any axis to quantify the specific characteristics of each. Furthermore, the number of emotions depends on the theoretical framework. Dimensional models define some common continuous features and place each emotion in a point in the space based on the values of the considered features. One of the most accepted dimensional models is the Russel Circumplex Model [302]. It has two dimensions: Arousal (degree of activation) and Valence (degree of pleasure). More complex dimensional models, in addition to Valence and Arousal, also include Dominance (degree of attention) to produce Valence Arousal-Dominance (VAD) or Pleasure-Arousal-Dominance (PAD), creating a 3D space of emotions [303]. Valence goes from unpleasant to pleasant, Arousal goes from passive to active, and Dominance goes from submissive to dominant, representing the degree of controllability of a specific emotion. Another model [304] is made up of Valence, Arousal, Dominance, and Predictability, where the last element represents the level of surprise. Among many, EEG is one of the techniques for measuring brain activity in real time and under almost normal, minimally invasive conditions. It has the advantage of having a great time resolution but a low spatial resolution and, compared to others, being cheaper and more portable. However, the accurate study of emotions would require their direct measurement. The measurement of spontaneous emotions in daily life conditions is challenging, due to many condunding variables that occur

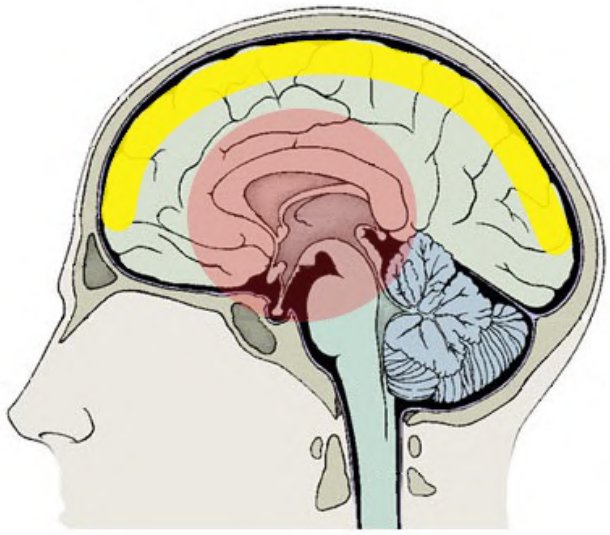


Figure 6.2: EEG brain sensitivity map (in yellow), with sensitivity decreasing with depth, and the main circuits of the limbic system, where emotions mostly originate (in red).

during the recording. Usually, for EEG, external stimulation (videos, sounds, images, etc.) or memory recall is required to elicit emotions in well-defined experimental sessions and environments. The resulting signals are usually stored in public datasets [305] to allow open source recovery, processing, and analysis. In particular, many of them use external stimuli such as videos or / and audio [97, 306, 307], and some of them use self-stimulation from their memory [308, 309]. Emotion recognition is a particularly difficult domain due to the number of variables influencing emotions, including the elicitation mode, the experimental protocol, the basal emotional state of the subject, and the instrumentation used. The main problem with data sets collected through external elicitation is the presence of biases, mainly related to the protocol used. For this reason, in recent years, many emotion signal data sets have been proposed, each with pros and cons, but no one is completely free from biases.

6.1.2 Elicitation protocols, EEG datasets, and challenges

An important aspect necessary to recognize emotions by EEG is to obtain specific characteristics that characterize their signals, collected from several analyzed subjects. This is done through the design of specific elicitation protocols and the organization of measurement sessions during which the elicited emotions are synchronized with the measurement, making it very difficult to measure spontaneous emotions in a real-life context. Protocols produce stimuli in the form of items, events, or conditions that cause a person to elicit emotional responses or behaviors to study and understand various psychological processes. Stimuli can include situations, scenarios, or social interactions. The well-known stimuli for eliciting the targeted emotions are Virtual Reality (VR), images, video games, music, audio/video clips, audio, and/or videos [305, 310–312]. Based on the type of stimulus, various emotions are elicited which are manually ranked to be used by nonlinear processing strategies. Without going into specific details about the elicitation protocols and the related datasets that originated from them, it is worth noting that several criticalities take origin from every stimulation protocol is possible to define. In general, significant challenges arise in obtaining reliable and representative EEG emotions data, with many limitations arising from elicitation approaches, lack of adequate baselines, reduced number of sensors, equipment instability, etc. In summary, it is possible to observe numerous challenges that must be overcome to obtain EEG data that more accurately represent emotions, with the goal of exceeding the current limitations caused by stimulation protocols, the lack of adequate baselines, the limited number of sensors during acquisition and instrument instability. The number of EEG channels should always be very high (> 64) to allow high-density acquisition, which can then be

reduced during preprocessing to identify the most significant channels, and a multimodal acquisition should always be used rather than a single system (e.g. EEG + Electrocardiography (ECG)). In addition, environmental parameters must be measured to verify and correct instrumental errors. User interaction with the experiment (e.g., evaluation of each stimulus) introduces numerous artifacts and interruptions that could introduce further bias and noise into the measured signal. Furthermore, a simple limit to overcome concerns acquiring all possible demographic, physical conditions (age, hair, health, etc.) and psychological data to control for potential confounding variables (e.g., lateralization) and using validated psychological scales for assessing emotional states. Additionally, using many different trials for each recorded emotion allows for intra-subject and inter-subject analyses, and the acquisition context should be as neutral as possible to avoid introducing additional biases due to the environment. Finally, the subject's adaptation time to the new experiment should be respected to allow them to feel comfortable, and consideration should be given to the subject's fatigue for long-duration experiments, including assessing their mental workload at another time. Among the current limitations specific for emotional studies, is found the use of individual emotional elicitation trials for each participant, as well as the significant subjectivity of responses to individual external emotional stimuli, which thus require prior and/or posterior evaluation to assess whether such stimuli are appropriate for the subject. Often, there is a lack of an adequate baseline in databases, which should instead be recorded at numerous points before, after, and during EEG data acquisition and also it is necessary to acquire their personal "ground truth" of the stimuli used, acquiring their evaluation of the stimuli used in the experiment. Finally, there is the ethical problem that must be taken into account during emotional stimulation. In emotional stimulation protocols, images or videos containing emotional situations are often used to elicit emotional reactions in participants that are commonly referred to external situations (e.g. people in critical situations) and not related to the participant's personal situation. It is important to note that this approach tends to focus mainly on negative emotions, as they cannot be induced directly in participants. In contrast, positive emotions can be evoked directly or through indirect stimulation showing that third parties are experiencing positive emotions. It is crucial to consider that, under normal conditions, participants are in a state of relaxation, which can be interpreted as a positive emotion according to many theories, including Russell's model, which posits that there are no "neutral" emotions [302]. Finally, the progress in using EEG to recover emotions depends on the ability to overcome the above issues effectively, when possible.

The main problem with datasets collected through external elicitation is the presence of biases, mainly related to the used protocol. For this reason, in the last years, many signal datasets of emotions have been proposed, each having pros and cons, but no one being completely free from biases. In what follows, a new protocol for the elicitation of emotions aiming at overcoming some of the biases affecting those previously proposed datasets, though maybe introducing its own is proposed to contribute to the advancement of understanding emotions. In the next section, an overview of the emotional datasets is provided and their strengths and limitations are discussed, thus preparing the way to the dataset proposed.

The datasets reviewed in this section, chronologically reported in Tab. 6.2 in some of their relevant characteristics, are public data sets that have been widely used in the last few years by the scientific community. A brief description is reported for all.

	IMAGINED EMOTIONS	**DEAP**	**MAHNOB-HCI**	**SEED**	**ASCERTAIN**	**DREAMER**	**SEED-IV**	**AMIGOS**	**SEED-V**	**MUSEC**	***REMEMO***
publication date	*2009*	*2011*	*2011*	*2015*	*2016*	*2017*	*2018*	*2018*	*2019*	*2021*	*2024*
n subjects	32 (13 M, 19 F)	32 (16 M, 16 F)	27 (11 M, 16 F)	15 (7 M, 8 F)	58	23 (14 M, 9 F)	44 (22 M, 22 F)	40	16 (6 M, 10 F)	20 (10 M, 10 F)	17
n. eeg channels	250	32	32	62	1	16	6	14	62	62	64
Acq. freq. (Hz)	256	512	256	1000	32	128	1000	128	1000	1200	1000
questionnaries				Eysenck Personality Questionnaire (EPQ)	Big-Five						Mood; PSQI; STAI; BDI; TAS; OHBDS; MCMI-III
stimulus		40 (60 s) music videos	first exp: 20 (34.9 s) second exp: 14 (117 s) movie videos and weather forecasts	15 (240 s) movie videos	36 (80 s) movie videos	18 (400 s) movie videos	72 (120 s) movie videos	20 (5-14 min) movie videos	9 (120-240 s) non specified videos	110 (19-66 s) song audios	text 70 (20 s) memories 35 emotional + 35 neutral
measured emotions	POSITIVE: Happiness, Relied, Love, Joy, Compassion, Excitement, Awe, Contentedness; NEGATIVE: Disgust, Sadness, Frustration, Fear, Grief, Anger, Jealousy; NEUTRAL: Relax;		Fear, Joy, Neutral, Amusement, Disgust, Sadness, Surprise, Anxiety, Anger	Positive; Neutral; Negative		Amusement, Excitement, Happiness, Calmness, Anger, Disgust, Fear, Sadness, Surprise	Happy; Sad; Neutral; Fear;	happiness, sadness, fear, neural, surprise, anger, disgust	Happy, Sad, Fear, Disgust, Neutral		happiness, sadness, anger, relax, disgust, fear, surprise + neutral memory
assessment		arousal, valence, dominance, liking, familiarity.	emotional tag, arousal, valence, dominance, predictability.	PANAS Scale: emotions experienced, prior exposure, understanding content.	arousal, valence, liking, familiarity, engagement.	arousal, valence, dominance.	valence, dominance. PANAS Scale: emotions experienced, prior exposure, understanding the content	emotional tag, arousal, valence, dominance, liking, familiarity.		arousal, valence, familiarity.	arousal; valence; dominance before and after the recording
preprocessing		CAR; downsample: 256 Hz highpass filter: 2 Hz		downsample: 256 Hz; filter (0.3 to 50 Hz); artifacts manually removed;		Artifact removed manually: eye blinking, eye movement, cardiac interferences by FIR filters, ASR; CAR.		highpass filter: 2 Hz Blind Source Separation: Entropy Rate Minimization	bandpass filter: 1 to 50 H; artifacts manually removed	highpass filter: 0.1 Hz; ICA; notch filter (50 Hz); CAR reref (2 - 45 Hz);	1-100 Hz; PREP pipeline (robust CAR, channel noise detection + interpolation, D.C. filter) ; ICA + IClabel
strength	audio stimulation before the task	fixation cross as baseline, multiple repetitions	short clip as neutral task, multiple repetitions	hint as baseline, multiple repetitions	fixation cross as baseline, multiple repetitions	neutral clip, multiple repetitions	hint as baseline, multiple repetitions	fixation cross as baseline in short video session, multiple repetitions	fixation cross as baseline, multiple repetitions	multiple repetition with and without assessment	absence of stimuli
weakness	single repetition, absence of baseline	audio and video stimulation during task, interruptions due to assessment	audio and video stimulation during task, interruptions due to assessment	audio and video stimulation during task, interruptions due to assessment	audio and video stimulation during task, interruptions due to assessment	audio and video stimulation during task, interruptions due to assessment	audio and video stimulation during task, interruptions due to assessment	audio and video stimulation during task, interruptions due to assessment. Absence of fixation cross in long video session.	audio and video stimulation during task, interruptions due to assessment	audio stimulation, absence of baseline	subjective memories, no assessment during recording
biases	single repetition	activating stimuli	activating stimuli	activating stimuli	activating stimuli	activating stimuli	activating stimuli	activating stimuli	activating stimuli	activating stimuli	memory can be difficult to evoke and with the same intensity

Table 6.2: Comparison between all datasets

6.1.2.1 IMAGINED EMOTIONS

The Imagined Emotion data set was created by the Swartz Center for Computational Neuroscience at the University of California San Diego [308]. The data set is made up of 32 healthy participants (19 women, 13 men, 18 to 38 years of age, mean and standard deviation: 25.5 ± 5 years old). A 250-channel high-density EEG system with EOG and ECG electrodes was used for data acquisition sampled at 256 Hz. The dataset was created to investigate the frequency modulators to determine the relationship of modulations in muscle and brain sources in EEG signals. Each participant was asked to imagine emotional scenarios for 16 different emotions. These emotions were positive-valence, negative-valence, and neutral and are shown in Tab. 6.3.

Category	Emotions
Positive Emotions	Happiness; Relief; Love; Joy; Compassion; Excitement; Awe; Contentedness
Negative Emotions	Disgust; Sadness; Frustration; Fear; Grief; Anger; Jealousy
Neutral Emotions	Relax

Table 6.3: Emotions analyzed in IMAGINED EMOTIONS dataset.

The reason for incorporating such a wide range of emotions was to oversample the space of similar emotions. In fact, for each participant, a single measurement session was performed for each imagined emotion. Participants were asked to explore the similarities between certain emotions (for example, "joy" and "happiness") and to assess the differences between emotions that share common characteristics. This approach deviates from the conventional orthogonal reference system for emotions. Each session started and ended with 2 minutes of closed eyes in a relaxed state, and each task began with a voice guiding for 15-30 sec the imagination of the scenario for the next 3-5 minutes. The authors used three different experimental protocols to record the feelings of the imagined emotions. The first group, made up of 16 subjects, communicates the onset of the emotion by pressing a button. The second group, composed of 11 subjects, pressed the button each time they felt the emotion, and the pressure-sensitive key indicates the intensity of the emotion. The third last group consisting of five subjects pressed the button when they felt a "surge" in emotion. The analysis performed by the authors showed that there were no differences between those three groups, so they merged these three groups into one single group. The duration of the emotions was between 43 seconds and 12 minutes (with a mean and standard deviation of 218 ± 94 s) and each experimental session was approximately 80 min for each participant. The limitations of this approach were the use of a single trial for each emotion and the requirement of manual interaction with the participants. In Fig. 6.3 the experimental protocol is shown.

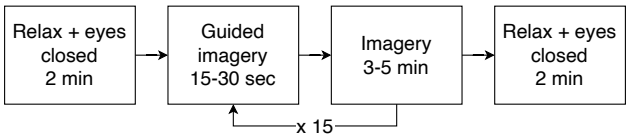


Figure 6.3: Experimental Protocol for the IMAGINED EMOTIONS dataset.

6.1.2.2 DEAP

The Dataset for Emotion Analysis using Physiological signals (DEAP) [97] is composed of EEG data from 32 subjects (16 women and 16 men) stimulated by 40 music videos, selected from an initial list of 120 by online raters. Each video was rated using different parameters: (1) arousal, (2) valence, (3) dominance, these dimensions on a discrete scale from 1 to 9, and (4) the emotion felt. Starting from 120 videos, based on the rating of many volunteers using a web-based subjective emotion assessment interface, 40 videos were

selected. From each video, 1 minute of highlights was selected using an algorithm [313] able to capture the most intense period of the video taking into account the audio and visual components, both for the arousal and valence components. The EEG used was an instrument with 32 channels sampled at 512 Hz, also using EOG channels, galvanic skin response, respiration belt, plethysmograph and temperature. Each subject watched 40 videos, and after each clip, the subject had to rate valence, arousal, dominance, liking, and familiarity. The first four dimensions had to be evaluated from 1 to 9 while the last, familiarity, was between 1 and 5. Finally, the DEAP dataset is made up of 32 people with 40 1-minutes of EEG signals recorded during external stimulation of the music video and 5 values per minute, with 5 seconds of baseline for each trial. One of the important aspects of the protocol DEAP is that for each subject 2 minutes of baseline were recorded in the experiment, beginning with a fixation cross on the screen, asking the participant to relax. Then 40 trials were presented with the following structure: 2 seconds during which the number of the current trial is recalled, 5 seconds for the fixation cross, 1 minute of the music video and the assessment with Self-Assessment Manikin (SAM) [314]. The EEG signal collected while fixing the cross on the screen, during relaxation, is used as a baseline signal, that is, the signal that has been eliminated from those allowing emotions. Then 40 trials with this structure were presented: 2 seconds indicating the number of trials, 5 seconds of the baseline with a fixation cross, 1 minute of the music video, and the assessment with Self-Assessment Manikin (SAM) in which the participant had to express arousal, valence, dominance, and liking on a scale of 1 to 9. SAM [314] is a non-verbal pictorial assessment method designed to directly evaluate pleasure, arousal, and dominance as they pertain to an individual's emotional response to various stimuli. This technique utilizes images of a cartoon character to depict different intensities of these dimensions, enabling participants to efficiently and effectively assess their emotional reactions. After the experiment, each participant expressed familiarity on a scale from 1 to 5. The dataset also includes the video recorded from a subset of 22 participants. The authors shared the preprocessed files for Matlab and Python on the DEAP dataset website[2]. It is important to note that the DEAP dataset has a huge subjective component: in fact, each subject can react differently to the view of the same video. This is the reason why the authors had to show the evaluation SAM during the trial. However, this could be another important issue because the subjects were forced to concentrate during the session to evaluate each trial during SAM, thus distracting them from the experiment. Another important source of biases is represented by the fact that both videos and audios were used to elicit emotion: brain activation due to videos and audios could obscure the emotional signal. Furthermore, the signal used as the baseline, that is, the EEG signal collected during the fixation cross, is not a proper baseline, since it lacks the audio component. In Fig. 6.4 the experimental protocol is sketched.

6.1.2.3 MAHNOB-HCI

The Multimodal Database for Affect Recognition and Implicit Tagging (MAHNOB-HCI) [306], is a database acquired during the showing of affective stimuli. The EEG signals were acquired during images and video exposure from 27 subjects (16 females and 11 males) of different cultural backgrounds, with 32 electrodes with 256 Hz of sampling rate, peripheral physiological signals such as Galvanic Skin Response (GSK), electrocardiogram (ECG), blood volume pulse (BVP) with plethysmograph, electrooculogram (EOG), respiration pattern, skin temperature at 60 Hz, the recording of audio signals (44.1 kHz), facial expressions with 6 camera (60 FPS), and eye gaze (60 Hz). The authors used two different experimental protocols: first with videos and second with images and videos. For video selection, 155 manually chosen video clips from 21 movies were presented to 50 subjects, and each clip received an average of 10 annotations using the algorithm presented in [313]. Online volunteers rated each video on the dimensions of valence and arousal using a Likert point scale from 1 to 9 with SAM and reported their emotions in response to vision. Finally, 14 videos from the online annotations were chosen (the most representative based on the number of emotional words), and the authors chose three other videos (one for "disgust" and two for "joy" emotion). In addition,

[2] https://www.eecs.qmul.ac.uk/mmv/datasets/deap/

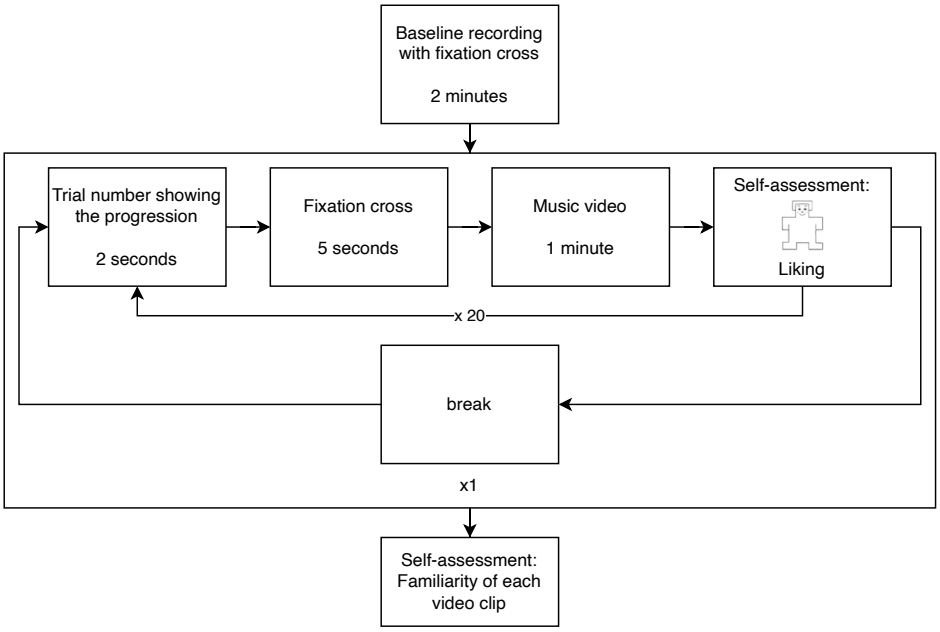


Figure 6.4: Experimental protocol for DEAP dataset.

three other videos of the weather forecast report were included and used as neutral stimulation clips. The duration of each clip was between 34.9 and 117 seconds (with an average of 81.4 seconds and a standard deviation of 22.5 seconds). The short duration of the video was justified for the exclusion of the interaction between different emotions that could be evoked by a longer video. The participants were exposed to images and video clips. After each video, they had to answer five questions, regarding the description of the felt emotion and its arousal-valence-dominance-predictability values (each ranging from 1 to 9).

Question	**Evaluation Type**
emotional label/tag:	one emotion among nine
arousal:	value from 1 to 9
valence:	value from 1 to 9
dominance:	value from 1 to 9
predictability:	value from 1 to 9

Table 6.4: Emotional Evaluation for the "Emotion recognition experiment" in MAHNOB-HCI.

For the first question, participants had to indicate the most representative emotional tag among nine predefined emotions, shown in Table 6.5. Each of the 20 videos was interleaved with a neutral video clip.

Fear	Joy	Neutral
Amusement	Disgust	Sadness
Surprise	Anxiety	Anger

Table 6.5: Emotions analyzed in MAHNOB-HCI database.

The neutral clip served to reduce the bias of the emotional state. After each clip, the participant answered five questions shown in the Tab. 6.4. In Fig. 6.5 the experimental protocol is shown.

In the second experiment, named "Implicit taggging experiment", 28 images and 14 video clips were shown to the participants. The videos were selected from the Hollywood Human Action Database (HOHA) [315] and the images from the Flickr website[3], accompanied by an incorrect or correct tag. For each trial, the

[3] `www.flickr.com`

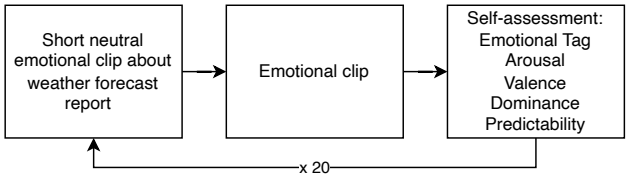


Figure 6.5: Experimental Protocol for the "Emotion recognition experiment" of MAHNOB dataset.

stimulus was first displayed without a tag and, after 5 seconds, with a tag. Then a question was displayed asking the participant if they agreed with the tag or not, pressing a green or red button. in Fig. 6.6 the protocol of the second experiment for MAHNOB-HCI dataset is shown.

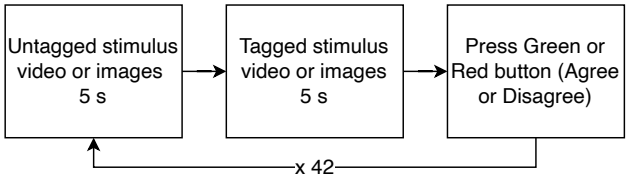


Figure 6.6: Experimental Protocol for the "Implicit taggging experiment" of MAHNOB-HCI dataset.

As in the DEAP dataset, the MANHOB data set is also affected by potential biases due to the subjectivity of the effect due to stimulus and due to the necessary concentration on the evaluation of the trial during the experiment. As for DEAP, both audio and video brain areas are requested, though in this case real baseline video clips were used, since both audio and video were present.

6.1.2.4 SEED

In the SJTU emotion dataset EEG (SEED) [316], from 15 right-handed people (7 men and 8 women) the EEG signal was acquired with a 62-channel system at 1000 Hz while showing emotional video clips. Fifteen videos from six Chinese films were selected and divided into 3 emotional categories: positive, neutral, and negative. All participants were selected using the Eysenck Personality Questionnaire (EPQ) [316, 317]. Each video was about 4 minutes long; Before the clip, a hint was shown to start for 5 seconds, and after that there were 45 seconds of self-assessment and 15 seconds for rest. The self-assessment part contained three questions used to ascertain the felt emotion during the video, an above experience regarding the video, and the understanding degree of the seen video [318].

- Inquiring about the emotions they experienced while watching the film clip.
- Wondering if they had any prior exposure to this movie.
- Inquiring if they understood the content of the film clip.

In Fig. 6.7 the experimental protocol for the acquisition of the SEED dataset is shown.

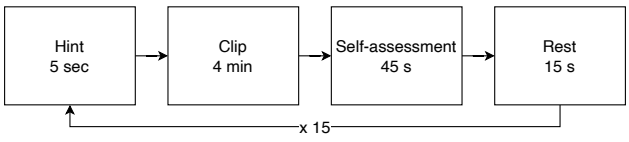


Figure 6.7: Experimental Protocol for the SEED dataset.

The critical aspects of the SEED dataset are the presence of both video and audio stimulation and the lack of a baseline signal. In addition, no left-handed subjects were monitored.

6.1.2.5 ASCERTAIN

The multimodal databaASe for impliCit pERsonaliTy and Affect recognitIoN [319] is made up of EEG signals collected from 58 people who watched movie clips and rated each clip with values of valence, arousal, engagement, liking and familiarity. The interesting characteristic of this dataset is that the authors, in addition to the EEG signals, also collected personality tests from all participants. The main difference between this dataset and others is the personality data, which used the five-big model that describes human personality on five dimensions:

- Extrovestion
- Neuroticism
- Agreeableness
- Concientiousness
- Openness

The set of video clips consisted of 36 videos from [320], with a duration between 51 and 128 seconds and an average of 80 seconds. Each participant watched 2 blocks with 18 videos for each block. In each trial, a fixation cross was shown before the clip and, at the end, the participants filled out the form with five questions about the feeling. The personality questionnaire was filled out by the participants at the end of the experiments. The complete protocol scheme in Fig. 6.8 is shown. Data acquisition also included ECG, GSR, and facial landmark trajectories (EMO). The limitations of this dataset were the usage of audio-video elicitation and the lack of a correct baseline signal, since just fixation cross was used.

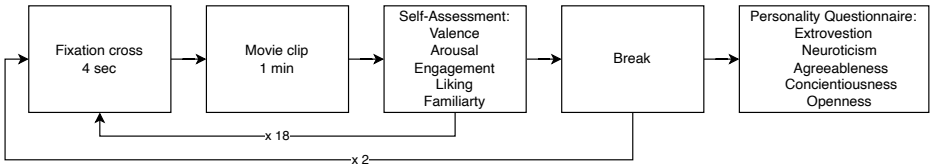


Figure 6.8: Experimental Protocol for the ASCERTAIN dataset.

6.1.2.6 DREAMER

The DREAMER dataset [307] is made up of EEG data from 23 people, 14 males and 9 female, between 22 and 33 years of age (average = 26.6, standard deviation = 2.7). The authors made this dataset to create a model that is useful in daily life conditions, with portable and non-expensive equipment EEG. Each participant was exposed to 18 video clips, two for each emotion, selected by [321]. The emotions selected are reported in Tab. 6.6.

Emotions		
Amusement	Excitement	Happiness
Calmness	Anger	Disgust
Fear	Sadness	Surprise

Table 6.6: Emotions analysed in DREAMER database.

The duration of each video was between 65 and 393 seconds with an average of 199 seconds, but only the recording of the last 60 seconds was kept to avoid multiple emotions. The EEG used was the Emotiv EPOC system [322], composed of 16 sensors in the 10-20 international system, sampling rate of 256 Hz. The session started with a neutral clip from [321] as a baseline, and a neutral clip was also shown after each emotional

clip. After each clip, the participant was asked to rate the clip using the SAM questions on a 5-point scale. The artifacts were removed using FIR filters, the Artifact Subspace Reconstruction (ASR) method [323] and CAR [103]. Limitations of DREAMER dataset were the usage of a very unstable EEG equipment with a very low number of sensors, the presence of artifacts and noise due to measurements in daily life conditions, the rating interaction required to the participants during the test. in Fig. 6.9 the experimental protocol is shown.

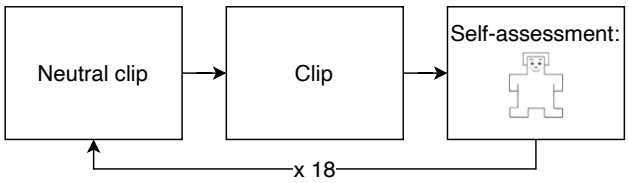


Figure 6.9: Experimental Protocol for the DREAMER dataset.

6.1.2.7 SEED-IV

With the SJTU emotion EEG dataset IV (SEED-IV) [324], the authors aimed to build a dataset composed by EEG and eye movement signals to improve the emotional recognition accuracy by reducing the artifacts due to eye movements. EEG data from six electrodes (FT7, FT8, T7, T8, TP7, TP8) of the 10-20 international system were collected and sampled at 1000 Hz. Eye traking was recorded during video stimulations, divided into 4 emotional categories, reported in Tab. 6.7. The authors performed the study using 168 video clips from 15 right-handed people (7 males and 8 females), age ranging from 20 to 24, and after each clip, each participant had to assess his valence (from sad to happy) and arousal (from calm to excited) on a scale from -5 to +5. After this preliminary study, 72 clips were selected based on video rating to create a set of the best clips for the emotion-evoking protocol. The duration of each clip was 2 minutes and there was no repetition of the same clip. The experiment was divided into three parts, each part consisting of 24 clips and six trials for each emotion. The execution of the experiment followed the scheme used for SEED and it shows almost the same limitations. Like in the previous version of the SEED, before each clip was shown 5 seconds of hint and at the end the PANAS scale [325] was shown for 45 seconds to assess the stimulation, based on the feeling of the viewer about the clip. If the subject certified that he did not feel any emotion, the data are discarded. The final dataset is composed of 15 people (7 males and 8 females) all right-handed, and their age ranged from 20 to 24 years old. in Fig. 6.10 the experimental protocol is shown.

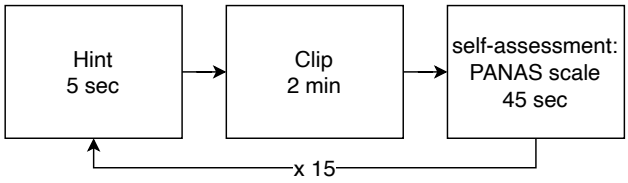


Figure 6.10: Experimental Protocol for the SEED-IV dataset.

Emotions	
Happy	Sad
Fear	Neutral

Table 6.7: Emotions in SEED-IV dataset

6.1.2.8 AMIGOS

The AMIGOS dataset [326] is composed of EEG signals acquired in two different contexts: the first is the "individual setting" with a series of short emotional videos (< 250 seconds of duration) and the second in the "group setting" with a series of long emotional videos (> 14 minutes of duration). Besides EEG signals, data from galvanic skin response (GSR), electrocardiogram (ECG), and RGBD videos with Microsoft Kinect V1 were also recorded. The authors hypothesize that in real life, affective experiences and social context play a crucial role in emotions. The reason to build this dataset is to integrate physiological information acquired using EEG and personality, mmood, and social group setting information. The EEG instrument was an Emotiv EPOC Neuroheadset3 with 14 channels in a 10-20 system sampled at 128 Hz. The participants were 40 healthy people between the ages of 21 and 40 with a mean of 28,3. All of them were self-assessed on the dimensions of arousal, valence, and dominance of the seven basic emotions (happiness, sadness, fear, neural, surprise, anger, and disgust) before the experiment. Then the set of 16 videos was shown in random order. Before each video, a 5 seconds fixation cross was shown to record the baseline. At the end of each video, the participant was invited to rate arousal, valence, dominiance, liking, familiarity, and emotional tag based on seven considered emotions. in Tab. 6.8.

Emotions
happiness; surprise; anger; disgust; sadness; fear; neural

Table 6.8: Emotions analyzed in AMIGOS dataset.

The participants in the two experimental sessions were the same. Video selection for the "short video setting" experiment (came from [320]) was carried out by 72 volunteers who annotated a set of 26 videos on the dimensions of valence and arousal, to place the videos on the Russell Circumplex Plane. Then, from each quadrant, the best three videos were chosen, for a total of twelve videos. Furthermore, an additional set of four videos, one for each quadrant, was taken from MAHNOB-HCI [306], for a total of 16 videos. The duration of the videos was between 51 and 150 seconds with an average of 86.7 seconds. For the "long video setting" experiment, starting from a set of eight videos taken from the IMDb Top Rated Movies list (`http://www.imdb.com/chart/top`), four researchers placed the videos on the Russell Circumplex Plane based on their value in the arousal and valence dimension; finally, the most representative subset of four videos, one for each quadrant, was selected. The lengths of this set are between 14.1 and 23.58 minutes, with an average of 20 minutes. In the "short video setting", 40 participants fill out a self-assessment form before the session and at the end of each video, one annotation was acquired for a total of 17 annotations based on 16 videos. For the "long video setting" 37 of the initial 40 participants were divided in two main groups: one made by 17 people is the 'individual session' and the other 20 people are grouped in 4 groups by five, matched by relationship between them or cultural background. In this second experiment, 6 annotations were acquired: one before the first subset of the first two videos, one before the second subset of the last two videos, and one after each video (four videos in total). in Fig. 6.11 the protocol of AMIGOS dataset protocol acquisition is shown. The data was filter with an high-pass filter at 2 Hz and cleaned with Entropy Rate Minimization, an blind source separation tecniques [327].

The interesting characteristic of the AMIGOS dataset is that it provides, besides EEG, also other physiological information. Limitations are that it used a very limited EEG equipment, with a very low number of sensors, with both audio-video stimulation, required active rating during the experiment, and did not measured a baseline.

6.1.2.9 SEED-V

In the SJTU Emotion EEG Dataset (SEED)-V [328], the EEG with 62-chanels system and eye movement data were recorded from 16 subjects (6 males and 10 females, age from 19 to 28), during emotion elicitation

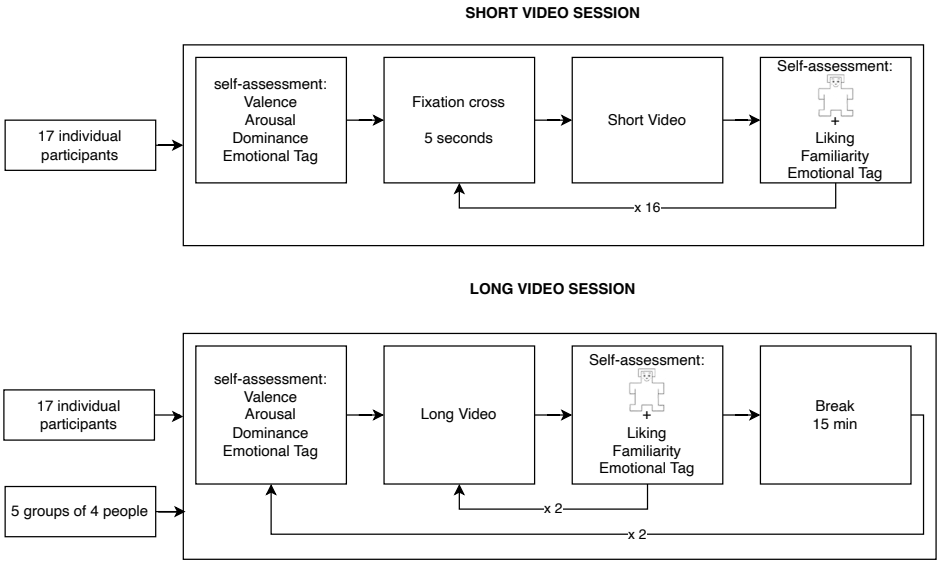


Figure 6.11: Experimental Protocol for the AMIGOS dataset.

using audio-video clips. The scenes were representative of five emotions: happy, sad, fear, disgust, neutral. The authors acquired this dataset to improve the stability, with respect to their previous work [316, 324], of the neural emotional pattern across three sessions using the same instruments used before (EEG and eye tracking), and investigating the critical frqeuency bands analyzed in [316]. For the preliminary study, 20 participants assessed the simuli rating each clip on a scale from 0 to 5. Then, 9 videos were selected based on the highest value of ratings, and each video had a duration from 2 to 4 minutes. All participants were part of three sessions of experiments in three times in different days, to investigate the stability of emotional patterns over time. For each session and for each emotion, three video clips were shown, for a total number of 15 trials. During the experiment, the sequence of opposite emotions such as happiness after sadness was avoided, but between them a neutral video clip was shown. Before the first session, each participant completed the Eysenck Personality Questionnaire (EPQ) [317]. in Fig. 6.12 the experimental protocol is shown. Besides the increment in stability, the SEED-V dataset has approximately the same limitations of

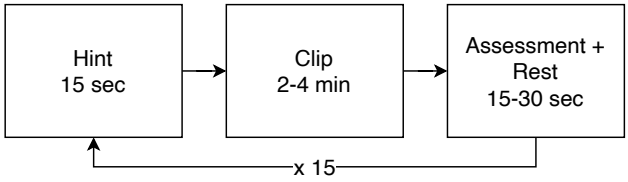


Figure 6.12: Experimental Protocol for the SEED-V dataset.

the previous versions of the same dataset.

6.1.2.10 MUSEC

The MUSEC data set [329] consists of EEG data from 20 participants equally divided by sex (age between 22 and 35) while listening 110 popular Thai musics selected from Spotify, divided in melodies and songs. The authors constructed this dataset to investigate the relationship between preference, familiarity, response rate and EEG information in response to music stimulation. In this dataset, data from 62-channel EEG and EOG sampled at 1200 Hz were recorded, being the length of each music between 19 and 66 seconds. The authors divided the music data set, of 110 songs, into two groups: melody (music without lyrics) and song (music with lyrics), . The melodies were played by an expert and converted to an MP3 file. To select the 110 songs from the initial set of 200, participants heard the songs in both modes and expressed their

preferences with a button. If both modalities were selected as favourites, the music and the melody were included in the set. The participants were divided into four groups of five people. For each group, $\frac{1}{5}$ of the 110 stimuli were randomly selected, and each participant heard 22 melodies first and after the 22 songs. For each group (melodies or songs) and for each stimulus, the participants had to click the play button; after the clicking, 5-seconds of silence were played to record the baseline; finally, a screen shown a message with the self-assessment questionnaire to rate the familiarity from -1 ('not at all') and +1 ('mostly'), and also the valence, the arousal and the 'adding to favorite playlist song.' Furthermore, the authors recorded the response rate to measure the velocity of the assessment to each music stimulus without informing the participants. First, the melody groups were played; after 5 minutes of break, the song groups were played. In Fig. 6.13 the experimental protocol is shown. The authors preprocessed data with a 50Hz notch filter, with a high-pass filter at 0.1 Hz and ICA artifacs correction. Finally, the data were referenced to Common Average Reference (CAR) a band-pass filter from 2 to 45 Hz were applied.

Limitations of the proposed dataset is the lack of a baseline, the necessity of rating during the experiment, and the subjective response to the stimulations.

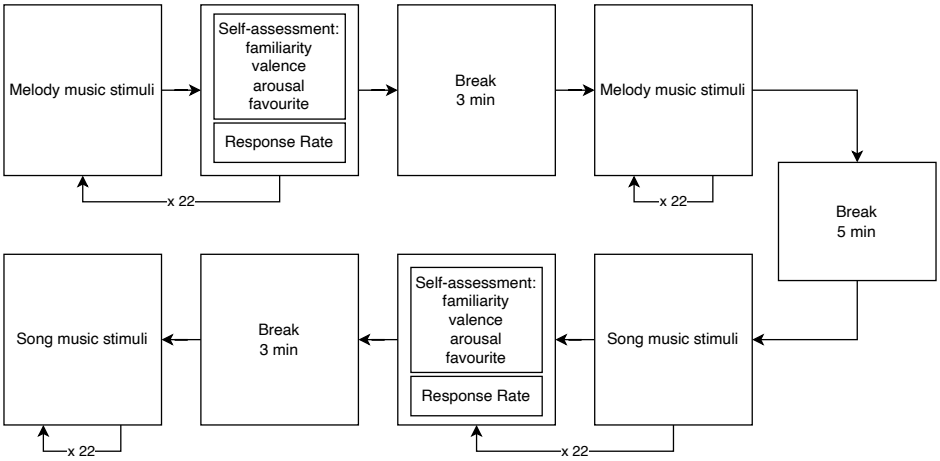


Figure 6.13: Experimental Protocol for the MUSEC dataset.

6.1.3 The surprising performance of the recent emotion recognition models

In the last few years, with the advent of nonlinear AI-based models, there has been a huge effort to improve emotion recognition by EEG [330, 331] and many more have been added since the above reviews. In particular, DL methods with CNN [332] and LSTM [333] have provided better performance, in terms of accuracy, than traditional deterministic processing algorithms. Indeed, over the years these models have been refined to include intrinsic characteristics of the problem, to get any potential spatial/temporal unknown relationship, and to better fit the data at hand. For example, the dynamic relationship among EEG channels in different regions has been represented graphically through dynamic graph CNN [334] and adaptive graphs [335], for taking into account the spatial correlations of EEG data. In addition, to include cross-subject variability and nonstationarity of the EEG signals, transfer learning has been introduced to solve the inconsistencies between training signals and the signals used for the test. To this end, new hypotheses were added to the models regarding brain symmetry / asymmetry, as in the adversarial neural network in the bi-hemisphere domain [336] and in the bi-hemispheric discrepancy model [337]. Other methods try to reduce subject variability in EEG by introducing specific domain residual networks [338, 339]. These models are similar to residual networks with the advantage that they do not require prior information during training. Several other model adaptations have considered the proposal of a dynamic adversarial adaptation network that dynamically learns domain-invariant representations both on a local and a global scale. One of these methods [339], proposed an adaptation of the joint distribution to take into account the differences between

individuals. Another recently proposed model [340] proposed an adversarial discriminative temporal convolutional network for the recognition of emotions in the domain of the other (subject and / or database) with further achievements in emotion recognition. For further details, several other methods for emotion recognition have been proposed, some of them are here reported [341–347], whose results have allowed to improve the accuracy that, for a three-class emotion dataset [348] (e.g. Positive, Negative and Neutral), ranged from about 60 %, from the first moves of DL, to above 90 % and, for a four-class emotion dataset [349] (e.g. Neutral, Sad, Fear, Happy), ranged from 38 % to above 80 %, with a huge improvement in the result stability (standard deviation passed from about 13 % of the first DL models, to about 7 % of the recently proposed models). Recent DL models applied to EEG have become increasingly complex in incorporating any kind of prior information and / or highlighting any sort of potential relationship among the measured data. In some cases, they also included limitations based on findings provided by other diagnostic tools, such as fMRI, to push the performance toward ideal values. Since most of the intrinsic mechanisms of brain behavior are still not completely understood, in particular regarding emotions, and because complementary measurement tools collect different parameters of the same phenomenon, consider different measurement brain places, are differently affected by noise and artifacts, and use different protocols, it could be improper to use the results of one to design constraints for the other. According to EEG, its main limitation in emotion recognition is that its measurements take place far from the source of emotions and that it probably measures the brain processing of emotions more than the emotions themselves. However, this does not mean that the measurements EEG are not useful for emotion recognition, but rather that the task could be difficult. In addition, the EEG signals are intrinsically nonlinear and nonstationary, and those from emotions are strongly subject-dependent and related to other brain processing tasks. Last but not least, emotions are not collected during real-life experience, but during the execution of a task elicited by a specific activation/synchronization protocol. Despite that, the performance of recent DL models for emotion recognition by EEG is surprisingly good, and doubt arises: The designed models could unintentionally use many heuristics, producing overfitting of unaware data. The present article aims to give rise to a reflection on the goodness of these results, to push, for the future, to a substantial verification of the proposed methods, to define further elicitation protocols to collect data to be used in this verification, to apply explainable AI to comprehend which features, brain regions and brain connections (spatial and temporal) are mostly involved in the emotional process and, finally, to apply these verified findings for EEG, emotion-driven, BCI.

6.2 Brain Functional Connectivity of Emotions

This section describes the brain connectivity analysis of EEG data from four self-stimulated emotional classes (“relax”, “anger”, “happiness”, “sadness”) using a graph model of PLI, which is PLI a measurement of connection insensitive to volume conduction effect. Qualitative results show that, for the analyzed emotions, connectivity analysis indicates some relevant differences both in the active brain regions and in the bandwidths involved in the activation. This method for connectome generation and analysis shows that useful information can be derived and used for contributing to disambiguating the problem of automatic emotion recognition.

Although classical analysis EEG evaluates each area of the brain separately with numerous metrics, such as PSD of individual channels, considering the interconnections among channels is also valuable [102]. Connectivity analysis, or connectome, is the study of information exchange between brain regions. The topological location of EEG channels can be usefully combined with PLI through graphs. The strength of using graphs applied to PLI of EEG is in the representation of brain interaction networks, which are fundamental to understanding the behavior of the network.

The authors in [102] analyze brain connectivity using three different indexes: Pearson Correlation Coefficient (PCC), Phase Locking Value (PLV), and Transfer Entropy (TE) on the DEAP dataset [97]. Furthermore, they use a data-driven approach to compute the connectivity matrix to maximize performance

using similarity and dissimilarity between EEG signals at different locations. After feature extraction, a CNN is used to classify emotional states reaching an accuracy of 80.73 % combining the PLV index with a distance equal to 5. For the emotion recognition problem also in [350] the DEAP dataset is used as input for CNN transformed into a set of connectivity matrices: PCC, PLV, and PLI. They also test three-order methods for the electrodes: one starting from the left frontal electrode (Fp1) and going over the head to the right frontal electrode (Fp2), first passing over the left hemisphere and after over the right. The second method consists of connecting all electrodes at a minimum distance without restrictions. In the third one, random order is used. The best classification result is 99.72 % using a CNN with 5 layers (3 convolutional layers and 2 max-pooling layers) with the second ordering method. For comparison, using PSD and the same CNN architecture, the accuracy is 80.86 %. Sun et al. [351] use Graph Theory applied to FC analysis in subjects with Major Depressive Disorder (MDD) compared to non-mdd subjects. They found that the imaginary part of coherence and cluster-span threshold unbiased binarization approach the right temporal brain region using Node Betweenness Centrality (NBC) in the alpha band, classify subjects with accuracy, sensitivity, and specificity on 87.50 %, 93.75 %, and 81.25 %, respectively. They use SVM as a classifier with the Radial Basis Function (RBF) kernel and the Leave-One-Subject-Out Cross-Validation (LOSOCV) validation method. They also perform statistical testing of the hypothesis with a p-value of 0.0001 with the permutation method. Peng et al. [352] use EEG network topologies to predict responses to antidepressant therapy in patients with MDD, using data collected before and after one week of antidepressant medication therapy. They analyze the brain networks of patients exploring brain fluctuation after one week and predicting efficacy for eight weeks prior. They find that the resting-state EEG network becomes significantly weaker after one week of treatment and that the resting-state network at baseline could be useful to distinguish responders compared to others with an accuracy of 96.67 % of spatial network topologies as features. Zangh et al. in [108] use PLI to compute the connectivity matrix and its binarization on Small-World Index (SWI). The result indicates alterations in frontal, temporal, and parietal-occipital regions of the left brain and temporal region of the right brain occur in MDD patients. Moreover, they find an average of the shortest path length, clustering coefficient in the left central region of theta band, and node betweenness centrality in the right parietal-occipital region are significantly correlated with the Patient Health Questionnaire-9 (PHQ-9) score of MDD. These three indexes are used as biomarkers to distinguish MDD from control with 93.31 % accuracy of classification. Liu et al. [104] use Graph Theory to analyze the functional connectivity of MDD patients during music perception. A decrease in connectivity pattern in the delta band and on the contrary an increase in connectivity pattern in the beta band compared to healthy people are discovered. Moreover, the healthy group shows lateralization of left hemispheric dominance compared to MDD people who don't show any lateralization. The classification with SVM of the data extracted by the beta band has obtained 89.7 %, 89.4 %, and 89.9 % of accuracy, sensitivity, and specificity, respectively. In [353], the authors aim to estimate EC using Generalized Partial Directed Coherence (GPDC) based on Granger Causality (GC) and used SVM to classify correctly left vs right-hand movements. The GPDC index determines the transient causal relation between different brain regions. Results show an accuracy of around 83.87 % in Alpha (α) band (8-12 Hz) and 83.05 % in the Beta-1 band (12-15 Hz). This section aims to use graph theory to analyze the brain wave pattern during self-stimulated emotions on EEG data contained in a dataset from the Swartz Center for Computational Neuroscience [308].

6.2.1 The Proposed Connectome Method

Graphs are mathematical models aimed at analyzing and representing relations by objectsas ordered pairs. in fact, a graph G is:

$$G = (V, E) \tag{6.1}$$

where each electrode is a node V and each functional connection between two electrodes is an edge E. Graphs can be applied for the representation and analysis of functional connectivity, providing useful information about the brain connectome in terms of topological properties. The index of each functional connection

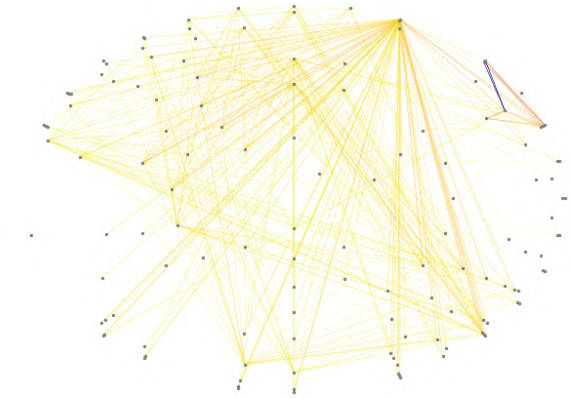


Figure 6.14: A graph of electrodes connection for β band of "happiness" emotion by subject n. 10. The strength of PLI for each pair of nodes is shown by the thickness and the color intensity of the line: higher thickness and more red represent an higher value of PLI. It can be noticed not all channels were included due to original authors exclusion of bad channels in the dataset.

is the strength of the edge, also called weight, *W(E)*. By applying this mathematical framework to brain connectivity, it is possible to compute several indexes [104]. In this section, PLI [106] is used as an index of the strength of each connection due to its insensitivity to volume conduction to compute the weights of the edges between the nodes. To compute PLI, the first step is to calculate, for each channel $x(t)$, the Hilbert Transform (HT), as in Eqs. 6.2 and 6.3:

$$z(t) = x(t) + i\widetilde{x}(t) \tag{6.2}$$

$$\widetilde{x}(t) = \frac{1}{\pi} PV \int_{-x}^{\infty} \frac{x(\tau)}{t-\tau} d\tau \tag{6.3}$$

where PV is the Cauchy Principal Value (CPV) and $\widetilde{x}(t)$ is the HT $x(t)$.

Then, the next step is to find the Instantaneous Phase (IP) of each channel:

$$\phi(t) = arctan \frac{\widetilde{x}(t)}{x(t)} \tag{6.4}$$

The third step is to calculate the difference of IP between two channels i and j:

$$\phi_{i,j}(t) = \phi_i(t) - \phi_j(t) \tag{6.5}$$

As a final step, the PLI is calculated (Eq. 6.6) between the two channels:

$$pli_{i,j} = |\frac{1}{T} \sum_{t=1}^{T} \text{sign}(\phi_{i,j}(t))| \tag{6.6}$$

where *sign* is the function defined as follows:

$$\text{sign}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases} \tag{6.7}$$

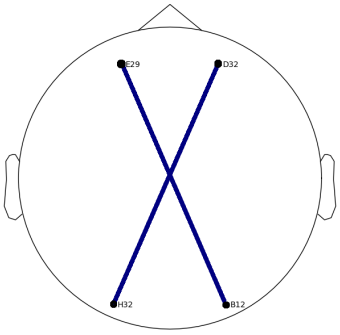


Figure 6.15: Simulation of PLI between four electrodes.

In Alg. 4 the whole process of connectome generation is shown, passing through the PLI computation.

Algorithm 4 Connectome Generator

for each frequency band k **do** ▷ Iterate over each frequency band
  for each channel i **do** ▷ Process each channel individually
    calculate HT of channel i ▷ Obtain the analytic signal to extract phase information
    calculate IP of channel i ▷ Compute the phase angle of the signal at each time point
  end for
  for each pair of channels (i, j) where $i \neq j$ **do** ▷ Compute pairwise measures between different channels
    calculate IP difference between channels i and j ▷ Determine phase difference between channel signals
    calculate PLI for channels i and j ▷ Evaluate phase synchrony between channel pairs
  end for
  construct graph $G(k)$ based on PLI values for band k ▷ Create a connectivity graph for the frequency band
end for
combine all graphs $G(k)$ into a single connectome object ▷ Aggregate graphs across frequency bands to form the connectome

PLI oscillates from 0 to 1. PLI is equal to 0 when two signals have no coupling or when the phase difference among them is 0 (in this last case, indicating that an extraneous source is generating the common coherence); PLI equals 1 when the two signals are completely phase locked. It should be noted that the distribution of the phase difference greater than or less than zero is symmetric [104]. PLI computation was performed using a modified version[4] of the EEGRAPH library for Python [112]. To check the validity of the method proposed in Alg.4, first was applied it to simulated data on the original set of 250 channels. To do this, signals from four channels, E29 (left frontal), B12 (right occipital), H32 (left occipital) and D32 (right frontal), were generated: E29 and B12 with the same frequency of 10 Hz and 10 μV and a constant phase lag of $\pi/3$; H32 and D32 also with the same phase and frequency of 25 Hz and 15 μV and a constant phase lag of $\pi/4$. These channels were chosen to better show the graph over the scalp considering their far-reaching positions, one for each quadrant of the scalp. An example of the graph is shown in Fig. 6.14. Then, PLI was computed for all combinations of these four electrodes, resulting in an indirectly connected dense graph weighted with four nodes and two edges, one for each couple. The graph reported in Figure 6.15 correctly shows the synchronization between channels with the same phase lag.

[4]https://github.com/danielelozzi/EEGRAPH

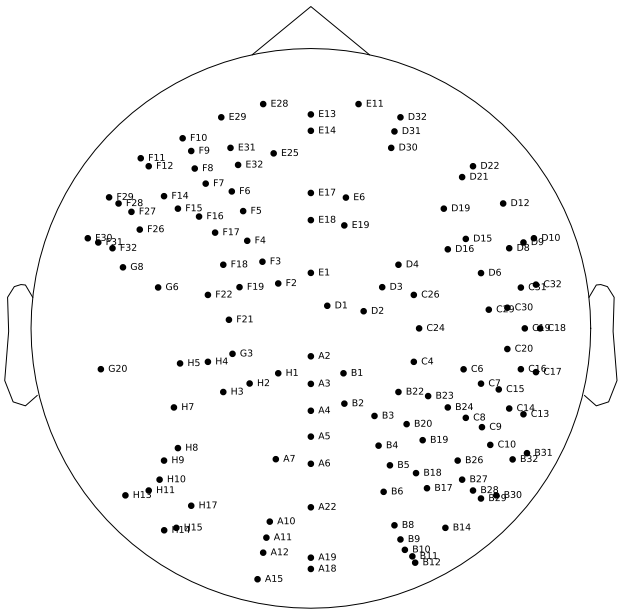


Figure 6.16: Positions of the residual common electrode (133 channels).

6.2.2 Experimental Results and Discussion

Once the method has been proven to be correct for simulated data, it has been applied to real data from emotions contained in the data set furnished by the Swartz Center for Computational Neuroscience [308]. The data set is made up of 32 healthy participants (19 women, 13 men, 18 to 38 years of age) sampled at 256 Hz, who were asked to imagine emotional situations for 16 emotions: positive-valence emotions ("happiness", "relief", "love", "joy", "compassion", "excitement", "awe", "contentedness"), negative-valence emotions ("disgust", "sadness", "frustration", "fear", "grief", "anger", "jealousy") and "relax". The reason for using so many emotions is to over-sample the space of similar emotions, asking the participants to experience similarities (e.g. "joy" and "happiness"), to check the distance among similar emotions. This approach differs from the standard orthogonal reference system for emotions [308]. In this preliminary study, only four emotions were analyzed: "sadness", "relax", "happiness", "anger". The rationale behind this choice was to select one emotion for each quadrant of the Russell Circumplex Plane. To create the signals, the authors of the dataset [308] asked to subjects to imagine an emotional scenario. Signals have been furnished preprocessed by an artifact removal system. In this section, data from five subjects (n. 1, 7, 10, 17, 27) were used, chosen by the high number of channels that these subjects share. Not all subjects have the same number of channels because the authors of the dataset [308] excluded very noisy channels, so only common channels between all subjects were left, for a total of 133 channels, as reported in Figure 6.16. This resulted in a not uniform covering of the head. The data was cleaned by applying band-pass filter from 1 Hz to 49 Hz to extract information physiologically relevant [354] and to exclude line noise. For each subject and for each emotion, the portion of the pre-processed signal from the first to the last marker was extracted and divided into frequency bands (δ, θ, α, β, γ) and with only 1-49 Hz band-pass filter. Then, 5 graphs of PLI, one for each frequency band, were computed and stored. The resulting calculation was composed, for each emotion, of five stacks of five graphs, one stack for each subject each consisting of five bands. For each emotion, the connectome graphs resulting from this computing divided by subjects and frequency bands are shown in Table 6.9 and the average by subjects in Table 6.10 (the average is also calculated for all subjects for each band, and is included in Table 6.9 on the row named *Avg*). The PLI index is coded by the thickness (higher thickness means a greater value of PLI) of the line and by its color, on *Plasma Reverse Palette*, also shown in the lower part of Tables 6.9 and 6.10. Before any discussion, it is important to note that the analyzed data are too small to allow a robust quantitative evaluation. However, though with just an attempted qualitative analysisa new paradigm

and new features to analyze complex data are presented, such as emotions. The resulting graphs shown in Table 6.9, for all the emotions considered except for ”sadness“, indicate a clear connection between the left anterior and the right posterior parts of the brain. This is mainly evident by considering the average in all the frequency bands shown in Table 6.9 on the last row of each emotion and in Table 6.10 where only the averages are reported. Another important aspect is the subjective response to emotions. In fact, given an emotion, a frequency drift is evident for different subjects. Moreover, a relevant amplitude difference in PLI occurs for different subjects. In detail: S. 17 shows very low levels of activation with respect to the other subjects, and this occurs for all the given emotions; S. 1 shows a similar pattern, though his activation is more marked in some bands (δ and α); S. 27, shows an approximately null activation to “relax”, a low activation to “anger” and an acceptable, though similar, activation to ”happiness“ and ”sadness“ (for S. 27, both the β and the γ bands are almost not used). This is also demonstrated by analyzing the numerical data in Table 6.11: this table shows the total sum of the PLI for each graph only with the 1-49 Hz band-pass filter. For all emotions analyzed, the pivotal band is the α band, as can be observed in Table 6.10. Regarding the specific emotions: ”anger“ concentrates on the α; ”happiness“ polarizes both on the θ and on the γ bands, with a decrease of α; ”sadness“ spreads to the whole frequency spectrum in almost the same way; ”relax“ polarizes to the left of the frequency bands (δ), and completely avoids the γ band. These considerations are reinforced by the data in Table 6.10. Table 6.12 reports the values of PLI, normalized to the maximum value to show the total number of weights summed in absolute values. As can be seen, the above considerations are confirmed by this synthetic table.

The proposed graph-based PLI representation method allows showing the connections between brain areas, improving the knowledge of how information relates to the brain, in particular for emotion analysis. From the graphs, some frequency bands seem strictly related to others, while some others seem not to be involved in the emotional process. In the same way, the proposed strategy allows for separating useful from useless frequency bands. Furthermore, this method could be in principle capable of showing, in a quantitative way, the common pattern across people or the subjective patterns of emotions and discovering the ”fingerprints“ of emotions. This fingerprint was founded for other tasks by several authors [355–357], but has not been investigated in the field of emotions yet. Future work will focus on collecting more experimental data to extend the study of EEG signals of emotions through graphs based on PLI, on structuring automatic procedures which exploit graphs of PLI, and on investigating the patterns of subjective connectome and fingerprints of emotions.

6.3 4D LSTM Neural Network for “relax” emotional state classification

Emotion measurement with EEG and their recognition through analytical models is based on the theory that emotions have specific patterns in cortical activation. The objective of this work is to build a DL architecture to discover these activation patterns from the PSD of the EEG signals from the public data set from the Swartz Center for Computational Neuroscience (University of California San Diego)[5] of self-induced emotions.

The emotion model used in this data set is in line with the Russel Circumplex Model [41], used during the data acquisition performed by the researcher at the University of California, San Diego. They used 15 emotions between positive and negative and “relax” before all self-inducted emotions, considered neutral. The authors justified the use of many emotions to over-sample the emotions and catch the similarity between

[5] https://sccn.ucsd.edu/

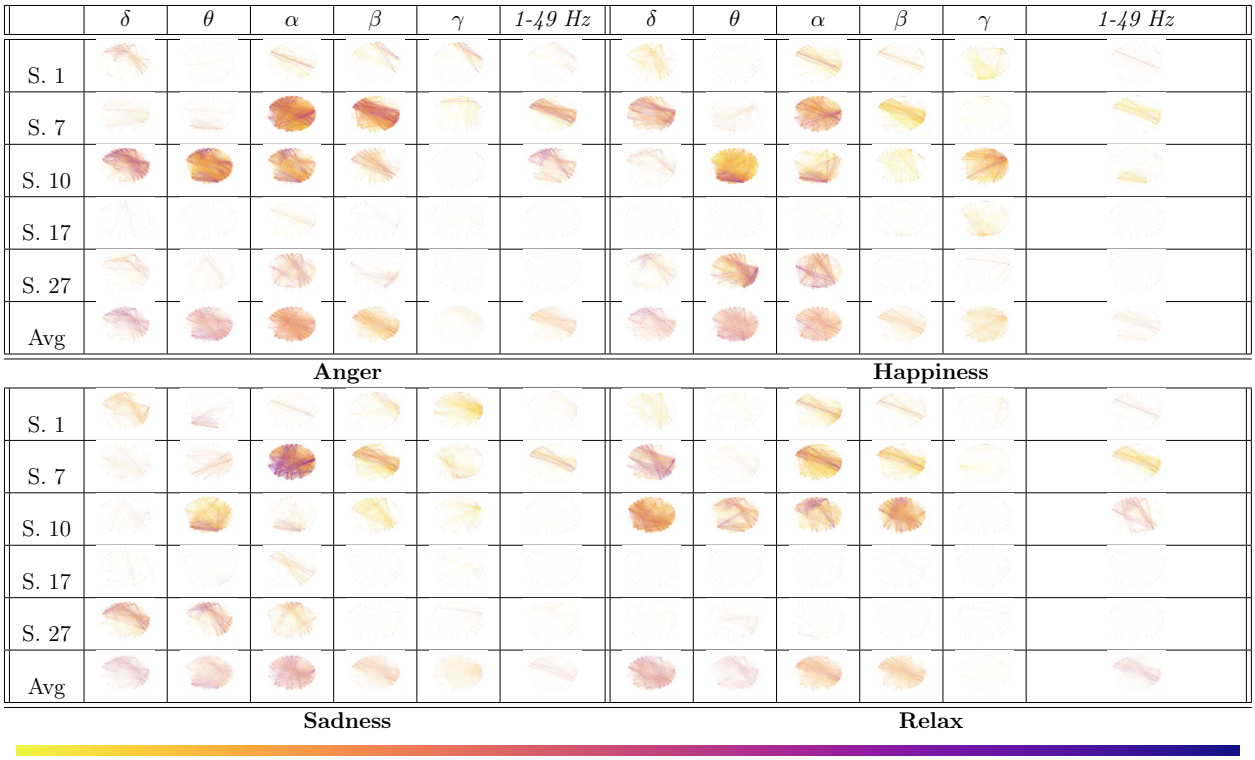


Table 6.9: Connectomes of four emotions located as in the Russel plane. For each emotion, the connectomes are divided by subjects and by frequency bands. The PLI value is coded by the thickness and the color of the edge. Under the table, color palette is reported. The PLI value goes from 0 (yellow) to 1 (blue).

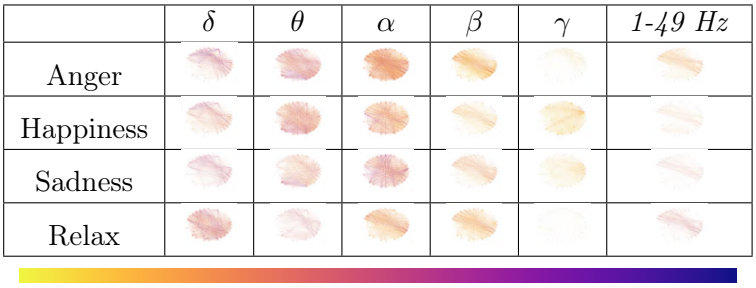


Table 6.10: Connectome averaged by subjects, separated for emotions and frequency bands. The value PLI is encoded by the thickness and the color of the edge. In the table, the color palette is reported. The value of PLI goes from 0 (yellow) to 1 (blue).

	S. 1	*S. 7*	*S. 10*	*S. 17*	*S. 27*	*Avg±St.Dev.*
Anger	17.8	168.9	125.1	0	0.1	62.4 ± 70.7
Happiness	13.1	68.9	38.2	0.2	1.4	24.2 ± 26.1
Sadness	11.5	73.7	1.9	0.0	8.6	19.1 ± 27.6
Relax	18.3	112.7	75.1	0.0	0.0	41.2 ± 45.1

Table 6.11: Sum of PLI for each subject filtered only with 1-49 Hz band-pass filter, separated for emotions and subjects

them. This is a different approach from the classical use of orthogonal emotions (like in many studies using the DEAP dataset), which classifies emotions only on valence and arousal. The choice of normalized cross-correlation it allows for testing the hypothesis that emotion has a specific pattern in terms of space (correlation of PSD over the head) and in terms of time (same pattern among correlation topoplot). The normalized cross-correlation could be useful because PSD change a lot during the development of people:

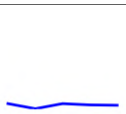
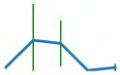
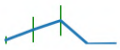
	S. 1	*S. 7*	*S. 10*	*S. 17*	*S. 27*	*Avg±St.Dev*
Anger						
Happiness						
Sadness						
Relax						

Table 6.12: Sum of PLI, separated by subject, by emotions and by frequency bands (δ, θ, α, β, γ are represented in the horizontal axis of each graph). Sum of PLI values are the vertical axis of each graph.

PSD is very different between adults and young people, so the normalized cross-correlation can represent very well the activity in terms of change over space and time, and it could be useful to compare different subjects with different μV in their power spectral activity.

In detail, for each subject, the EEG temporal signals was divided into partially overlapping "epochs" (segments of EEG signals with the same length based on experimental parading) and for each epoch the PSD was computed. The spectrogram was divided into the six spectrograms of the frequency bands of interest defined in the literature (delta, theta, alpha, sigma, beta, gamma). For each frequency band, the cross-correlation of each epoch with the following epochs was computed and related to their spatial position in the scalp.

A 4D LSTM network, based on 2D convolutional layers, is used to recognize emotional states, using spatial information (EEG channel positions), also by exploiting the temporal information (evolution over time) of the correlation among epochs. Preliminary results have been obtained by recognizing the "disgust" and the "relax", both intra- and inter-subject. The results show the convergence of the network but the accuracy of prediction on subjects varies greatly from subject to subject, in agreement with the most recent literature. Moreover, though intra-subject accuracy was sufficiently high for some subjects, the intersubject accuracy was very low, thus also demonstrating the subjective nature of the emotional patterns. Although a deeper analysis is required, this section presents a new model for the processing of emotional signals EEG.

Among these several interesting papers, all proposed the analysis of the signal aims to a specific instance as a single object, with rare exceptions [137, 275] in which the signal of an instance is broken into several sub-instances composed of partially overlapping windows. If, from one side, the consideration of the whole instance could be an advantage because it gives the possibility of integrating the signal along its duration (thus increasing the SNR), from the other side, this could be a limitation if the non-stationary nature of the neural signals was taken into account, in particular those related to emotions. Among the strategies reported, none has been found to use the cross-correlation of PSD and its evolution over time as input to the neural network to classify emotions. In order to investigate the possibility of analyzing emotional signals along time, a 4D LSTM network based on the analysis of the correlation of the PSD among channels (spatial), frequency bands and time is presented. The reason for using PSD is to be robust to phase changes in the signals when using temporal sliding windows. To our knowledge, this represents a completely new way of analyzing the signals EEG from emotions.

6.3.1 Materials and methods

To evaluate the proposed framework, data collected from the Swartz Center for Computational Neuroscience [358] were used and analyzed with the MNE-Python package for Python [177]. The data set is composed of 32 healthy participants (19 women, 13 men, from 18 to 38 years old) who were asked to imagine emotional situations for 16 emotions: positive-valence emotions ("happiness", "relief", "love", "joy", "compassion", "excitement", "awe", "contentedness"), negative-valence emotions ("disgust", "sadness", "frustration", "fear", "grief", "anger", "jealousy") and "relax". The reason for using so many emotions is to over-sample the space of similar emotions, asking the participants to experience similarities (e.g. "joy" and "happiness"), to check the distance among similar emotions. This approach differs from the standard orthogonal reference system for emotions [358]. Data have been preprocessed, that is, the data have been subjected to artifact removal and ICA to clean up the signals. A 256-channel high-density EEG system was used for data acquisition. In this preliminary study, was chosen to analyze only the emotion of "disgust" using "relax" as a baseline to perform a binary classification. For data augmentation, 7 sets of 32 electrodes were extracted from the original 250 (EOG and ECG electrodes were not used). Based on a 32-channel 10-20 system, for each subject, the following technique was used: for each iteration, the 32 electrodes closer to the 10-20 system configuration were used. For the next iteration, these electrodes were excluded. The procedure continued until there were less than 32 electrodes left. The signal was filtered using a bandpass filter from 1 to 45 Hz and downsampled it from 256 to 128 Hz. To create the signals, the authors of the data set [358] asked the subjects to imagine an emotional scenario. The participants indicated the start and end of each emotion by using a button, and these time windows were extracted. From the whole signal, 8-second windows (epochs) were extracted, each overlapping with the following by 4-seconds. It should be noted that, in this section, an *epoch* represents a specific time window extracted from the EEG signal and not the number of iterations in a data set. For each epoch, the PSD with Welch's method was computed to obtain the spectrogram. Subsequently, the spectrogram was filtered into six subspectrograms of specific frequency bands (delta, theta, alpha, sigma, beta, gamma). Then, for each channel and each frequency band, the normalized cross-correlation was computed between two following epochs and used as a feature between two spectrograms of different epochs of the same emotion.

The number of emotions recorded was one for each subject and the number of epochs (time window EEG signal) for each emotion depended on the time each subject experienced a specific emotion. After feature extraction, 383 4D tensors "relax" and 200 4D tensors "disgust" for training/validation were used, using the 80 % for training and 20 % for validation. The test was performed on every subject (see Tab. 6.13) using individual data each time. Cross-correlation is a measure of similarity between two discrete-time series. The normalization is made by computing the sliding dot product over the square of the product of the autocorrelation of every single spectrogram:

$$z[k] = \frac{\sum_{l=0}^{\|x\|-1} x_l y*_{l-k+N-1}}{\sqrt{(x \cdot x)(y \cdot y)}} \tag{6.8}$$

where x is the first time series and y is the second time series.

$$k = 0, 1, 2... \|x\| + \|y\| - 2 \tag{6.9}$$

$$\|x\| \text{ is the lenght of x} \tag{6.10}$$

$$\|y\| \text{ is the lenght of y} \tag{6.11}$$

$$N = max(\|x\|, \|y\|) \tag{6.12}$$

To obtain a cross-correlation map, the array was reshaped at the original channel position. Later, the points between the channels were interpolated in order to create a correlation map, for each frequency band, and used to prepare a 3D tensor (Fig. 6.17).

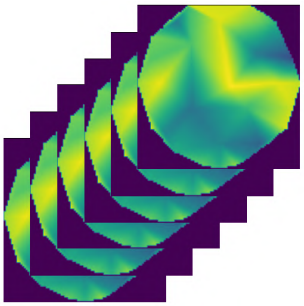


Figure 6.17: 3D tensor: correlation maps of one epoch generated for the six frequency bands.

Finally, the set of correlation maps (equal to *n-1* epochs) was used to create sets of temporal maps in sequence, to obtain for each channel the time evolution of the correlation (Fig. 6.18). To summarize, each sample was a 4D tensor composed of:

$$time * height_map * length_map * band \tag{6.13}$$

The whole process of 4D tensor production is summarized in Fig. 6.19

6.3.2 4D LSTM Neural Network

To analyze the 4D produced tensor, a 4D LSTM network was used, based on 2D convolutional layers, to classify emotions by capturing, in addition to spatial relationships between channels in different frequency bands, also the invariant information along time referred to the same emotion, as sketched in Fig. 6.20. This is done by combining 2D convolutional layers to extract the spatial relations, with the LSTM to extract temporal information, as reported in Fig. 6.21. The choice of a ConvLSTM layer instead of two different layers is to take advantage of both information in space (correlation among channels) and in time (correlation among time flow).

In particular, the used network is composed of 16 layers: 3 recurrent 2D convolutional layers, 4 dropout layers, 2 average pool normalization and 1 max pool normalization layer, 3 batch normalization layers, and one dense layer with 32 neurons before the classification stage. The activation function of the classification layer is the sigmoid.

The proposed framework has been trained, validated and tested on data coming from 28 of 32 subjects. Data were extracted considering two emotions: "relax" (baseline before emotions induction) and imagined

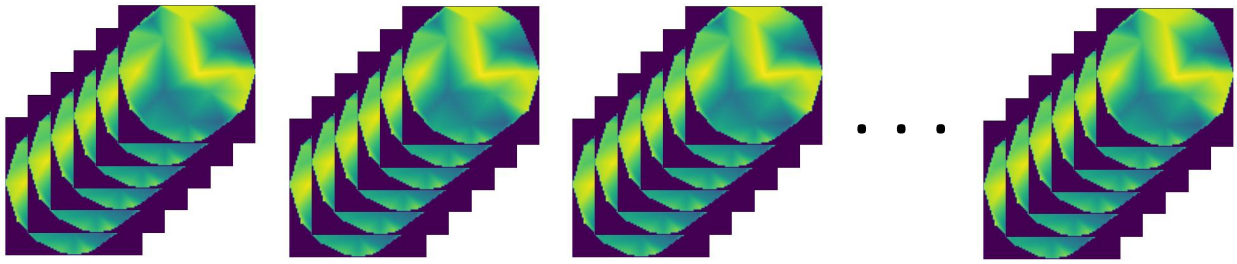


Figure 6.18: 4D tensor: temporal sequence of 3D tensors. The resulting 4D tensor is used as the input of the LSTM.

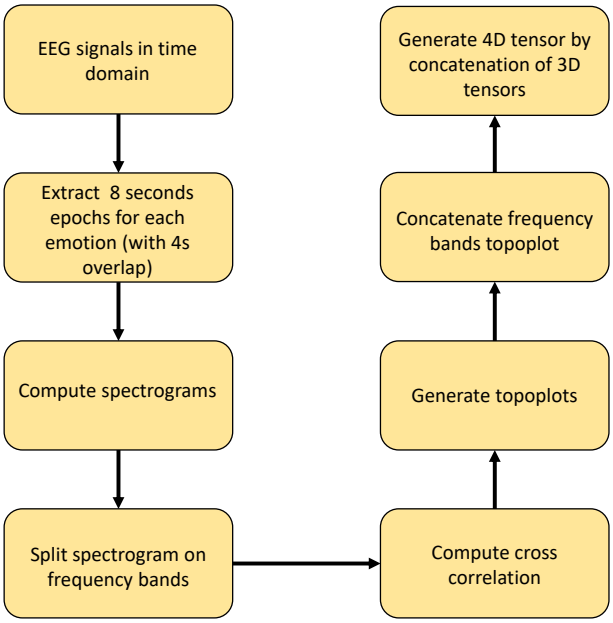


Figure 6.19: 4D tensor production flow-chart.

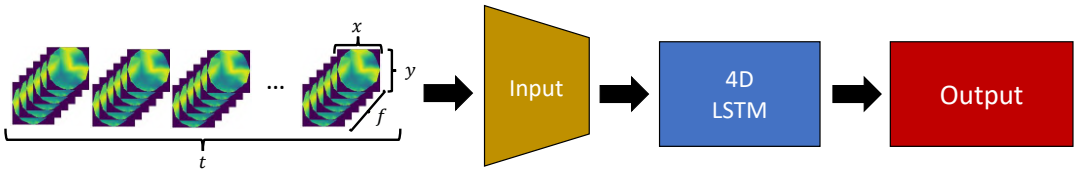


Figure 6.20: General scheme of the used architecture. In particular, the 4D tensor input and the binary classification output are reported. The internal design of the LSTM network is reported in Fig. 6.21

scene with "disgust" activation. Subject n. 31 was discarded because of its extremely noisy data. Subjects n. 3, 22 and 30 were not present in the data set). Data from subjects n. 27, 28, 29 and 32 were left out of training and validation and used just for testing the framework. This approach allowed for testing the proposed framework on completely unknown data. Data from all the other subjects were used for training, validation and testing, paying attention to ensure that data balancing was maintained among subjects used for training and validation. The accuracy and loss of the model, both during training and validation, are reported in Figs. 6.22 and 6.23, which indicates on the x-axis the number of iterations in the data set. The unusual behavior of accuracy and loss during validation can be explained by a high individual difference in pattern activation and to the low SNR in the EEG signal.

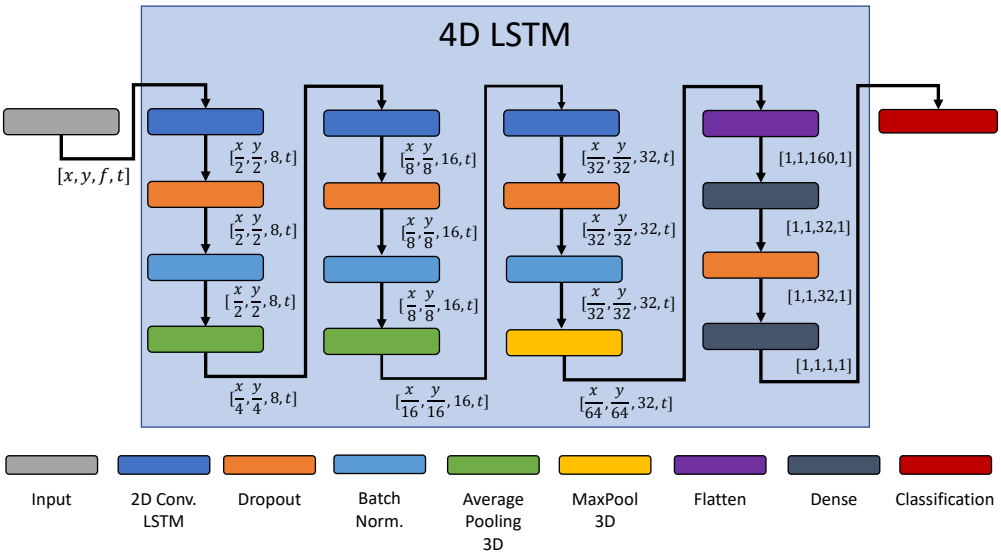


Figure 6.21: LSTM architecture in details.

The results obtained for the classification of "relax" and "disgust", separated by subject, are reported in Tab. 6.13. Subjects n. 27, 28, 29, and 32 were not included in the training and validation sets in order to use them to test the proposed architecture with data completely unknown to the network. The prediction was computed using only the first set of 32 electrodes, i.e. the most close set of 256 electrodes configuration to the 32 electrodes configuration. A simple test was performed on the 4 subjects who were left out of training and validation. Future experiments will consider LOSOCV to better evaluate the model used.

Our results, though still very preliminary, seem to indicate that an acceptable relationship in activation

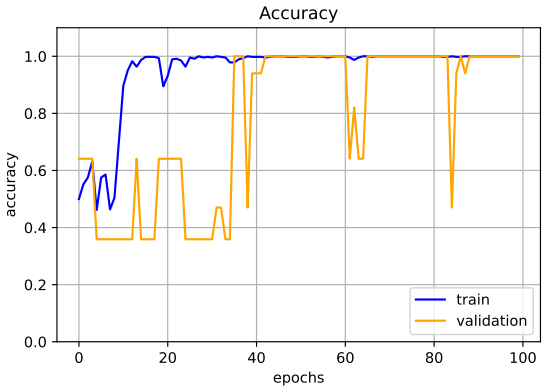


Figure 6.22: Model *accuracy* during training (blue line) and validation (orange line).

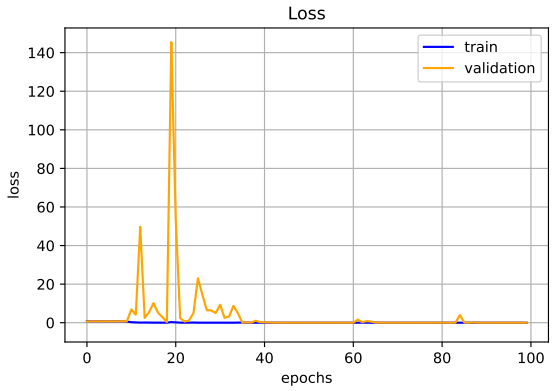


Figure 6.23: Model *loss* during training (blue line) and validation (orange line).

along time exists just for a very few subjects or, alternately, that there is a very low correlation among data from a different subject, though derived from the same emotion. This last consideration appears to be demonstrated by the fact that the results obtained for the last 4 subjects do not differ from chance. In this last case, results seem to confirm that, though there could be a temporal relationship among intrasubject data, an intersubject relationship appears to be completely absent. In this case, it is also justified why the framework performance is so bad when data from several subjects is used for training and validation. Furthermore, the results appear to indicate subjective behavior when activating a specific brain path for a given emotion.

Another crucial aspect could be represented by the fact that herein were compared the "disgust" with "relax": since there could be limited pathways when imagining the "disgust", conversely there could exist a very huge number of activation patterns for imagining the "relax", so in this section of binary classification, it is used "disgust" as activated emotion versus "relax" with the meaning of no activation [275]. Moreover, the choice of "relax" as the "reference emotion" is related to the data set construction: the authors of the data set recorded "relax" before all self-induced emotional trials. While a comparison with other studies would be beneficial, it is not possible at this time because, to our best knowledge, there does not exist any other study using this data set with the same aim.

Subj. n.	Accuracy	Subj. n.	Accuracy	Subj. n.	Accuracy
1	0,52	12	0,86	23	0,47
2	0,61	13	0,77	24	0,42
4	0,59	14	0,67	25	0,58
5	0,69	15	0,85	26	0,48
6	0,75	16	0,78	27	0,65
7	0,71	17	0,71	28	0,60
8	0,76	18	0,47	29	0,62
9	0,45	19	0,54	32	0,70
10	0,58	20	0,73		
11	0,48	21	0,67		

Table 6.13: Accuracy values, separated by subject, obtained during the test when recognizing "relax" and "disgust". Data from the 4 subjects shown in the gray background were not used for training and validation.

Emotion recognition, from EEG signals, is a complex and still open problem. In the present paper, a new framework based on a 4D LSTM neural network was presented, using 4D LSTM neural network for the analysis of the correlation between PSDs, at different frequency bands, of different channels and at different times. The proposed framework contains both 2D convolutional layers, for discovering spatial relationships,

and LSTM layers for registering temporal correlations. The binary classification results, obtained to distinguish "relax" from "disgust" in 28 subjects, were not particularly satisfying, except for a low number of them (9 subjects obtained an accuracy above 70 %). From one side, this could mean that there is a very low relationship between data from different subjects, though experiencing the same emotion, and, from the other side, this could mean that the proposed framework does not perform very well for this specific purpose. This section shows the innovative use of normalized cross-correlation as a feature extracted by PSD and the use of different architecture. Moreover, this method allows the architecture to be independent both of the number of channels and of the signal length (just PSD was used). Finally, this model can be easily extended to other emotions, by adding more data without changing the main architecture. The limitations of this section regard the quantity and the diversity of the utilized data. First, the data set used does not include repetitive acquisitions of self-induced emotions, something that would be useful for extracting useful information from noisy data. Furthermore, the number of subjects is reduced to 32, since a lot of criteria must be met for the inclusion of a participant. The participants were all healthy, while in a future study it would be interesting to include subjects with certain diseases and neurological conditions, to expand the study to people suffering from brain illness conditions such as depressive disorders. A future data set should have also long "relax" instances. Moreover, novel techniques could be used to perform artifact removal from EEG signals using AI-based methods [83]. Finally, future work should be also devoted to ascertaining the good behaviour of the framework.

6.4 Neural correlations of visual emotional processing

The analysis of ERP has gained significant attention in studies examining brain responses to images with positive and negative valence. Early studies in this field revealed that ERP components, such as P300 and N400, are particularly sensitive to emotional and cognitive stimuli, showing significant variations in response to images with different affective valences [359]. The P300 component, for example, has been associated with attention and evaluation processes, while the N400 has been linked to semantic processing, highlighting how emotions influence cognitive processing [359]. Recent studies indicate that the frequency bands involved in ERP responses primarily include the theta (4-8 Hz) and alpha (8-12 Hz) bands, which have been associated with memory and attention processes [360]. Moreover, EEG topographic analyses have shown that responses to negative valence stimuli tend to activate different brain areas compared to positive stimuli, suggesting differential involvement of neural networks. Negative valence has often been associated with increased activity in frontal and temporal regions, while positive valence tends to engage parietal areas [360]. Some studies have indicated that EEG responses are more intense for negative images compared to positive and neutral images, analyzing the P200 Event-Related Potential (P200) and P300 components [361]. The processing of neutral images requires less cognitive effort, as they do not carry the same emotional weight as emotionally charged images, resulting in smaller amplitude in the LPP, which is sensitive to the emotional content of the stimulus [362, 363]. This leads to a slower and less pronounced component compared to negative images. Since the frontal areas are highly sensitive to the emotional content of stimuli, analyzing electrodes F3, F4, and other frontal electrodes makes it possible to identify the P300, which is more pronounced for negative images [361]. The central channels, on the other hand, exhibit the P200, which is associated with cognitive control and increased cognitive load, and ERSP analysis should indicate specific oscillatory activity in the theta and alpha frequencies [364, 365]. The channels over the parietal areas are highly sensitive to the LPP, which should be more pronounced following emotionally charged images compared to neutral ones [366], in the beta and gamma bands [362]. Finally, the channels positioned over the occipital lobe should show a marked N170 Event-Related Potential (N170), indicative of visual processing. Regarding relevant electrodes, studies have employed high-density electrode configurations to achieve more precise mapping of brain responses [367]. The use of standardization techniques such as the Reference Electrode Standardization Technique (REST) has improved the quality of EEG data, enabling better identification

of ERP components and their variations in response to emotional stimuli. Additionally, PCA has been utilized to examine EEG activity in relation to task performance, revealing significant correlations between frequency bands and cognitive performance [360]. In one study [368], the authors examined how stimulus characteristics, particularly affective valence and visual properties (color versus black-and-white, normal images versus scrambled images), influence electrophysiological responses measured through the ERP P300. The results of the study indicate that normal color images generate valence effects on P300, suggesting that the affective valence of visual stimuli plays a significant role in cognitive processing. In particular, when arousal level is controlled, normal images produce a more pronounced P300 compared to scrambled images, which fail to evoke the same affective response. This suggests that normal images activate stronger perceptual and cognitive processing mechanisms than scrambled images, which, being altered, cannot convey the same affective valence. Moreover, the study highlighted that P300 is influenced not only by the affective valence of the stimuli but also by their visual characteristics. Color images were found to evoke a more significant P300 compared to black-and-white images, indicating that color plays a crucial role in visual information processing and emotional response. This aligns with existing literature suggesting that color can influence the perception and interpretation of images, contributing to a more intense emotional response. This study provides evidence that normal and color images generate stronger P300 responses compared to scrambled and black-and-white images, emphasizing the importance of affective valence and visual characteristics in modulating brain responses to visual stimuli. In [369], the authors explore how stylized faces, characterized by exaggerated visual traits, influence electrophysiological responses measured through ERP in emotion detection. The authors argue that using stylized faces may enhance the ability to detect emotions compared to normal faces, due to greater activation of specific ERP components, such as N170, known for its role in facial recognition. The study's findings indicate that stylized faces produce a more pronounced N170 compared to normal faces. This suggests that the exaggerated visual features of stylized faces facilitate the processing of emotional information, making emotion recognition more efficient. In particular, the N170 has been associated with facial encoding processes, and the increased amplitude in response to stylized faces indicates greater attention and deeper processing of the emotions conveyed by these stimuli. Additionally, the study discusses the importance of stimulus transformations in emotion research. Stylized faces, due to their simplicity and visual clarity, can reduce background noise and improve emotion detection, thereby facilitating the analysis of brain responses. This approach could have significant implications for the design of future experiments in psychology and neuroscience, particularly concerning emotion processing and facial recognition. In summary, this study highlights how stylized faces can enhance ERP features used for detecting emotional responses, suggesting that the use of simplified visual stimuli may be an effective strategy for studying brain dynamics associated with emotions.

In [370], the authors address the important topic of how lighting conditions influence emotion classification through DL approaches and neural analysis. The authors focus on the interaction between ambient light conditions and neural responses associated with emotions, using advanced ML techniques to improve accuracy in emotion classification. The study's context is based on the idea that human emotions can be influenced by environmental variables, including lighting conditions. Variations in light can alter not only visual perception but also emotional responses and associated brain activities. Therefore, the authors designed an experiment to examine how different lighting conditions might affect the accuracy of emotion classification using EEG signals. In their approach, the authors collected EEG data from participants exposed to various lighting conditions (e.g., natural light, artificial light, and darkness) while emotional stimuli were presented. Using DL techniques such as CNN, they analyzed the EEG data to identify neural features associated with different emotions, such as joy, "sadness", and "fear". The results showed that lighting conditions significantly influence the performance of emotion classification models, with some conditions leading to higher accuracy than others. Specifically, the study revealed that natural lighting conditions tend to improve emotion classification compared to artificial ones. This may be attributed to the greater authenticity and naturalness of stimuli presented in well-lit environments, which can facilitate emotional processing and neural responses. Additionally, the authors discussed the importance of considering environmental variables in emotion studies,

suggesting that integrating factors such as lighting could enhance understanding of emotional dynamics and brain responses. This results underline the importance of lighting conditions in emotion classification through EEG analysis and DL. The findings suggest that environmental variables must be carefully considered in emotion research, as they can significantly influence neural responses and their interpretation.

Based on the current scientific literature, the analysis of ERPs in response to positive and negative emotional images reveals distinct patterns of brain activity associated with different frequencies and cortical areas. ERPs are measures of the brain's electrical activity elicited by specific stimuli and are instrumental in understanding how emotions are processed at the neural level. As extensively discussed in Subsection 2.2.3, exposure to positive images generally leads to an increase in the amplitude of the LPP, an ERP component that reflects attention and emotional processing. This increase is particularly pronounced in the parietal and occipital areas, where the LPP reaches its peak [371]. Conversely, negative images can also evoke a LPP, but the response may vary depending on individual factors [372]. Additionally, negative images have been observed to activate components such as N1 Event-Related Potential (ERP N1) and P2 Event-Related Potential (ERP P2), which are associated with heightened attention and an immediate emotional response [373].

Regarding active brain frequencies, studies have demonstrated that positive emotions are associated with increased activity in the left frontal areas, whereas negative emotions tend to correlate with higher activation in the right frontal areas. This frontal asymmetry has been interpreted as indicative of an emotion processing model, where positive emotions are linked to stronger activation in the left hemisphere, and negative emotions in the right hemisphere [374]. Additionally, parietal and occipital areas exhibit different responses to emotional stimuli. Positive images tend to generate more intense activation in parietal regions, while negative images may strongly activate occipital areas, suggesting that visual evaluation and emotion processing differ depending on the emotional valence of the stimulus [371]. Finally, it is important to note that the neural response to emotional stimuli is not uniform and may vary due to factors such as age, mental health, and personal experiences. For instance, younger adults generally display a larger LPP compared to older adults, indicating that age can influence emotional reactivity and neural processing [375]. The ERPs computed in this section, shown distinct neural reactions to positive and negative images, with notable variations in brain wave amplitude and location. Positive emotions are linked with increased activity in the left frontal and parietal areas, whereas negative emotions elicit heightened activation in the right frontal and occipital regions. Furthermore, PSD has been used to study the brain's response to emotionally valenced images. In particular, in [376], the authors recorded participants' EEG signals while they were exposed to images from the International Affective Picture System (IAPS), including positive, negative, and "neutral" images. They observed a decrease in alpha power and an increase in gamma power in posterior areas in response to unpleasant images compared to pleasant ones (and also compared to "neutral" images for gamma power). The data analysis shown an increase in PSD in the gamma band in response to unpleasant images compared to pleasant and "neutral" ones, suggesting greater attention towards negative stimuli. This observation is corroborated by the reduction in alpha rhythm, an index of an increase in attentional cognitive load [377]. Indeed, scientific literature associates alpha rhythm with an inhibition of cortical activity and, consequently, with less attention. It is interesting to note how [377] reports an increase in gamma activity in response to emotionally salient images, both pleasant and unpleasant, compared to "neutral" ones.

As already illustrated in Chapter 3, an ideal BCI is composed of both an *active* and a *passive* system: the *active* component allows the recognition of a command that the user voluntarily sends to the system through their CNS, while the *passive* component reads the mental state of the subject following stimulation. The first system is therefore activated by an internal state, the second is instead induced externally [134]. In the context of Affective Computing, to build an ideal Emotional BCI it is necessary to structure a system that contains both types calibrated to recognize self-induced emotions (*active*) and emotions aroused by external stimuli (*passive*). To pursue this goal, a new dataset was presented whose objective is to structure an *active* BCI system, in which the user, through the stimulation of self-induced emotions from memories, can guide a BCI. In this section, the experimental protocol allows the acquisition of EEG data with three mental states,

in particular "happiness", "sadness" and "neutral", in order to compare the EEG signals acquired with the two experimental paradigms and in future studies build an ideal BCI, as shown in Fig. 3.28. The objective of this section is therefore to build a new dataset and carry out preliminary analyses to verify if there is information in the acquired EEG data that is consistent with the scientific literature by carrying out analyses on ERP, PSD and other features widely illustrated in Chapter 3. In details, the protocol that will be performed aims to understand neural correlates under the visualization of images from the Nencki Affective Picture System Basic Emotion (NAPS-BE) dataset [310, 378] and Set of Fear Inducing Pictures (SFIP) dataset [379].

6.4.1 Image datasets

6.4.1.1 Nencki Affective Picture System (NAPS) and Nencki Affective Picture System Basic Emotion (NAPS-BE)

The Nencki Affective Picture System (NAPS) dataset is a fundamental resource for emotion research, designed to provide standardized images that evoke affective responses. This dataset was developed to address the limitations of other affective image systems, such as IAPS, and to offer a wide range of visual stimuli covering diverse emotions and cultural contexts [310]. The Nencki Affective Picture System (NAPS) includes an extensive collection of images that have been carefully selected and evaluated for their affective properties. The images are categorized based on variables such as valence (positive or negative) and arousal (level of emotional activation), enabling researchers to utilize the dataset for studies on how emotions influence cognitive and behavioral processes [310]. Additionally, the NAPS was designed for use in cross-cultural research contexts, as demonstrated in a study comparing emotional responses in Europe and Iran [380].

The NAPS dataset is composed by 1356 high quality pictures divided in five categories:

- People
- Faces
- Animals
- Objects
- Landscapes

and each of them are labeled by 204 people (mostly from European countries). The labels are shown in Tab. 6.14 and in Fig. 6.24.

Valence	Arousal	Approach-Avoidance
Negative	Low	Avoidance
Positive	High	Approach

Table 6.14: Three dimension model used in Nencki Affective Picture System Basic Emotion (NAPS-BE) labeling process.

Participants were asked to rate their perception of the image on the valence scale by completing the sentence, *"You are judging this image as ..."* using a scale of 1 to 9, where 1 indicated a very negative perception and 9 indicated a very positive perception, with 5 indicating a "neutral" perception. Next, participants were asked to assess their motivational direction in response to the image by completing the sentence, *"My reaction to this image is..."* using a scale of 1 to 9, where 1 indicated a tendency to avoid and 9 indicated a tendency to approach, with 5 indicating a "neutral" reaction. Finally, participants were asked to evaluate the level of arousal elicited by the images using the introductory sentence, *"Confronted with this image, you are feeling:..."* using a scale of 1 to 9, where 1 indicated a relaxed state and 9 indicated a highly aroused state, with 5 indicating a "neutral" or ambivalent state. The rating were performed using a semantic bipolar scales and not with SAM due to the risk of misinterpretations. Furthermore, the authors provide users with the physical data regarding luminance, contrast, and entropy.

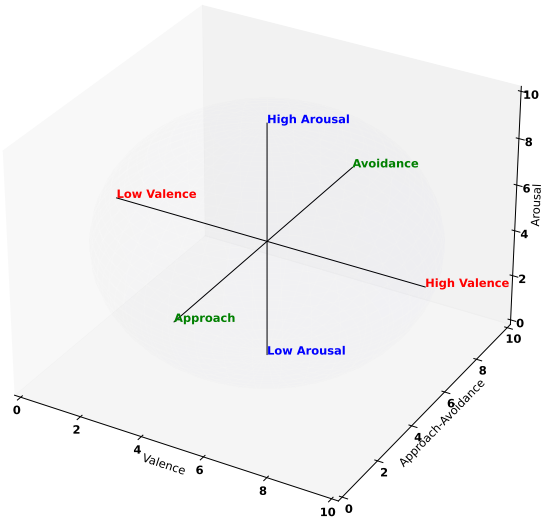


Figure 6.24: Dimensional space of the NAPS labels.

An extension of the NAPS is the Nencki Affective Picture System Basic Emotion (NAPS-BE), which focuses on the classification of basic emotions [378]. This subset of the dataset comprises 510 images annotated with six discrete emotions: "happiness", "fear", "sadness", "surprise", "anger", and "disgust". The images in the NAPS-BE have been evaluated to provide average scores for each emotion, allowing for a more fine-grained classification of emotional responses [381]. This segmentation of basic emotions is particularly useful for studies aiming to explore correlations between specific emotions and their neural manifestations, facilitating the analysis of brain responses to affective stimuli [381]. For this study, the classification based on "Briesemeister Liberal" parameters [382] and provided in the NAPS-BE materials were used because it consist in one cateogory of the six previously listed emotions for each images of the NAPS-BE dataset, selecting only "happiness" and "sadness" images from this dataset. The image selection algorithm applied is shown in Alg. 5.

6.4.1.2 Set of Fear Inducing Pictures (SFIP)

The Set of Fear Inducing Pictures (SFIP) dataset represents a thoroughly established and standardized repository of photographic stimuli that have been specifically assembled for the purpose of examining a range of phobic disorders. This dataset has been developed and validated through two distinct and independent research studies. It consists of a robust collection of 956 high-definition photographs, which are systematically classified into categories pertaining to social anxiety, arachnophobia (phobias associated with spiders), and the "fear" of blood/injections, along with a set of "neutral" images for comprehensive analysis [379]. The images within the Set of Fear Inducing Pictures (SFIP) set were rated by individuals experiencing social anxiety, spider phobia, blood/injection phobia, as well as by control groups without such "fear". These assessments took into account both the psychological implications of the samples and the intrinsic physical properties of the images, including parameters such as luminance, contrast, and overall complexity. In the context of this particular study, exclusively "neutral" images were extracted from the SFIP dataset. The selection criterion was based on the ratings provided by the control participants, with the selection procedure being guided by the algorithm outlined in Alg. 5. This approach ensured a rigorous and methodical process

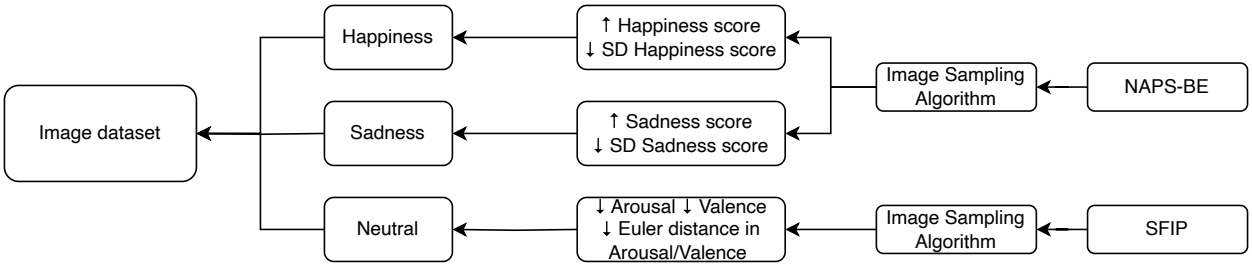


Figure 6.25: The image selection process

for image selection tailored to address the study's specific research objectives.

6.4.2 Image selection

The Alg. 5 is designed to select images from two datasets, Nencki Affective Picture System Basic Emotion (NAPS-BE) and Set of Fear Inducing Pictures (SFIP), based on specific criteria related to the emotions expressed in the images. The goal is to create three balanced datasets of images ("happiness", "sadness", and "neutral"). The algorithm selects and organizes images labeled "happiness" and "sadness" from the NAPS-BE dataset and the "neutral" from SFIP to create a balanced dataset suitable for this experimental study. For happiness images, the happiness ratings and their standard deviations are scaled to a range between -3 and 3, normalizing the data. Subsequently, images with scaled values less than or equal to 0 are eliminated, keeping only those that clearly represent a strong expression of happiness. The remaining images are sorted according to the scaled standard deviation, so as to favor those with a more stable emotional expression. From the ordered list, the first 60 images are selected. The same logic is applied to the sadness images: also in this case, the images are filtered, ordered and selected according to the defined criteria to ensure stability and consistency of the emotions represented. As for the neutral images, those labeled as "neutral" and coming from the SFIP dataset are considered. Before proceeding, all images already selected for happiness and sadness are excluded, eliminating duplicates. Valence and arousal ratings, along with their standard deviations, are normalized in the same range of -3 to 3. Subsequently, only images with scaled valence and arousal within the range [-1.2, 1.2] are kept, ensuring that they represent a neutral emotional state. At this point, an index called $global_SD$ is calculated, based on the standard deviations of valence and arousal, which represents the stability of the neutral expression. By ordering the images based on this index, the first 60 are selected. The images selected for the three experimental conditions are then merged into a single dataset. This dataset is divided into three equal parts, each containing the same number of images for each emotion. To ensure random presentation of images in an experimental setting, each dataset is shuffled, preventing more than three consecutive images from belonging to the same emotion. This process ensures that the final dataset is balanced, composed of images selected according to precise criteria. Fig. 6.25 shows the image selection process.

6.4.3 The experimental paradigm

Participants, recruited on a voluntary basis, were provided with a detailed description of the experimental protocol, including the objectives and procedures of the study. After signing the informed consent form, each participant was assigned a unique identification code to ensure anonymity throughout the data collection and processing stages. Data collection consisted of two phases: the first conducted in the laboratory and the second conducted online from home, with specific instructions regarding the environment.

The initial phase of the study involved administering a computerized questionnaire via the platform `Psytoolkit.org`[6], aimed at collecting demographic information (age, educational level, disciplinary train-

[6]Psytoolkit.org

Algorithm 5 Image selection Algorithm from NAPS-BE and SFIP

Load Datasets
Load NAPS-BE dataset.
Load SFIP dataset.
for images labeled with "happiness" from NAPS-BE **do**
    Scale "happiness" ratings to a range of -3 to 3.
    Scale standard deviation of "happiness" ratings.
    Filter out images with scaled "happiness" ratings ≤ 0.
    Sort by scaled standard deviation of "happiness" (ascending).
    Select the first 60 images.
end for
for images labeled with "sadness" from NAPS-BE **do**
    Scale "sadness" ratings to a range of -3 to 3.
    Scale standard deviation of "sadness" ratings.
    Filter out images with scaled "sadness" ratings ≤ 0.
    Sort by scaled standard deviation of "sadness" (ascending).
    Select the first 60 images.
end for
for images labeled with "neutral" from SFIP **do**
    Exclude images used for "happiness" or "sadness".
    Scale valence and arousal ratings to a range of -3 to 3.
    Scale standard deviation of valence and arousal ratings.
    Filter images with scaled valence and arousal within a specific range (-1.2 to 1.2).
    Calculate a $global_SD$ based on valence and arousal standard deviations.
    Sort by $global_SD$ and select the first 60 images.
end for
Create and Shuffle Datasets
Concatenate all selected images.
Create three empty datasets
Split the concatenated dataset into three datasets with an equal number of images from each category.
Shuffle each dataset until there are no more than 3 consecutive values of the same image type.

ing area, gender, and place of residence) and assessing the participants' psycho-emotional state using the following psychometric scales, described in detail in Appendix B:

- Positive and Negative Affect Schedule-eXpanded (PANAS-X);
- Arousal and Valence on a 1 to 9 Likert scale;
- Brief Mood Introspection Scale (BMIS);
- Mood on a -10 to 10 Likert scale.

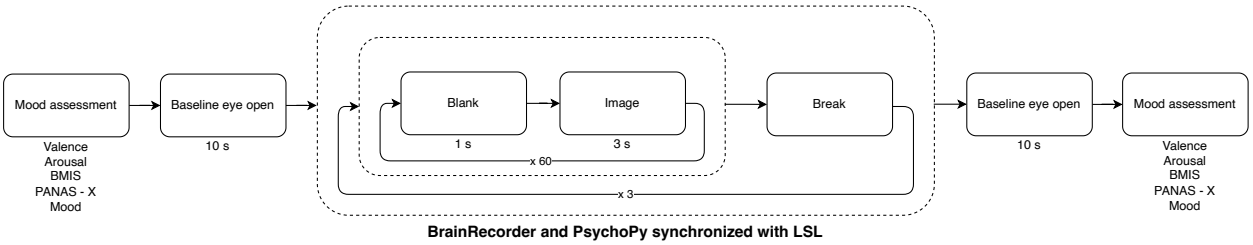
Subsequently, participants were subjected to an experimental paradigm involving the presentation of visual stimuli, during which EEG activity was simultaneously recorded. The paradigm, implemented using the PsychoPy software [383], consisted of the serial presentation of three blocks of stimuli, each comprising 60 emotionally valenced images (happiness, sadness, and neutral), as detailed in Section 6.4.2. The order of stimulus presentation within each block was randomized to avoid sequence effects and minimize habituation, with the constraint of not presenting more than three consecutive images from the same emotional category. Each stimulus was presented for a duration of 3 seconds, preceded by a 1-second Inter-Trial Interval (ITI) consisting of a black screen. Between blocks, participants were allowed to take self-managed breaks to mitigate fatigue. EEG signal acquisition was conducted using a 32-channel LiveAmp system based on the 10-20 system with wireless transmission technology [7] at 500 Hz, integrated with a trigger box for experimental event synchronization and Lab Streaming Layer (LSL), a software for real-time data stream synchronization and management [8]. The EEG acquisition protocol included a 10-second baseline period with eyes open before the paradigm's onset, followed by the presentation of the three stimulus blocks and an additional 10-second recording period with eyes open after the stimuli presentation. At the end of the experimental session, participants completed a second paper-based questionnaire, analogous to the first, to assess potential changes in psycho-emotional state induced by the experimental paradigm. The experimental paradigm is illustrated in Fig. 6.26 (a), while the recording environment is shown in Fig. 6.26 (b).

After 72 hours the experimental session, participants received an email inviting them to participate in the second phase, which included a link to access the second part of the study conducted online via the platform Psytoolkit.org [384, 385]. This phase involved administering a set of questionnaires:

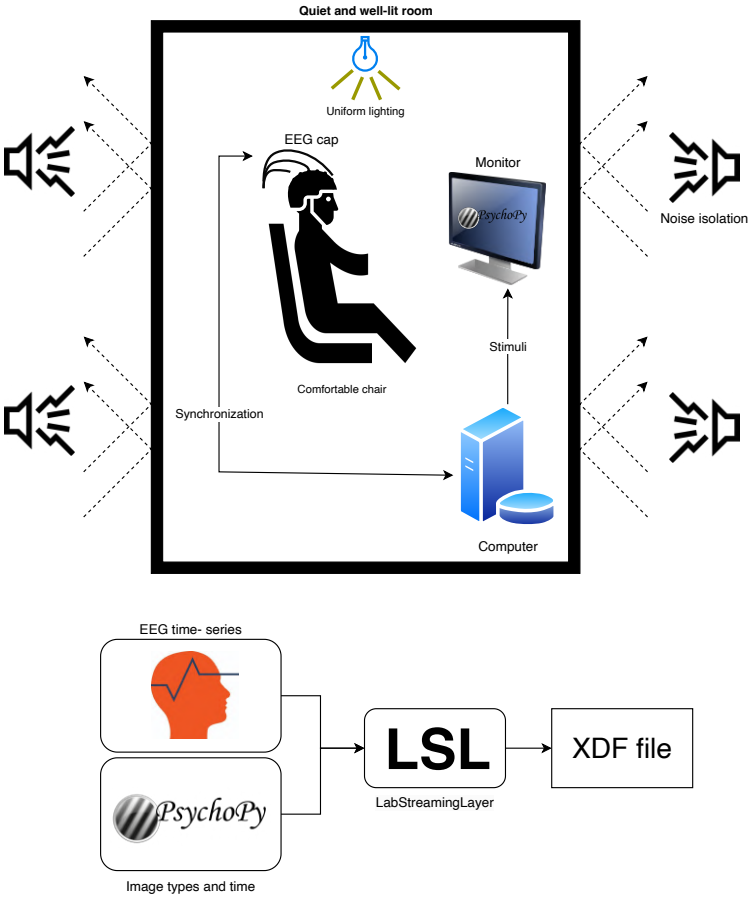
- Demographic data;
- Handedness question;
- Beck Depression Inventory (BDI)
- State-Trait Anxiety Inventory (STAI)
- Toronto Alexithymia Scale (TAS-20)
- Open Hemispheric Brain Dominance Scale (OHBDS)
- PANAS-X
- Arousal and Valence on a 1 to 9 Likert scale;
- BMIS;
- Mood on a -10 to 10 Likert scale.

Subsequently, participants completed an online experimental paradigm structurally analogous to the one presented in the laboratory, but without EEG acquisition and with a different order of visual stimuli presentation. After viewing each image, participants were asked to rate the intensity of the following

[7] https://www.brainproducts.com/solutions/liveamp/
[8] https://github.com/sccn/labstreaminglayer

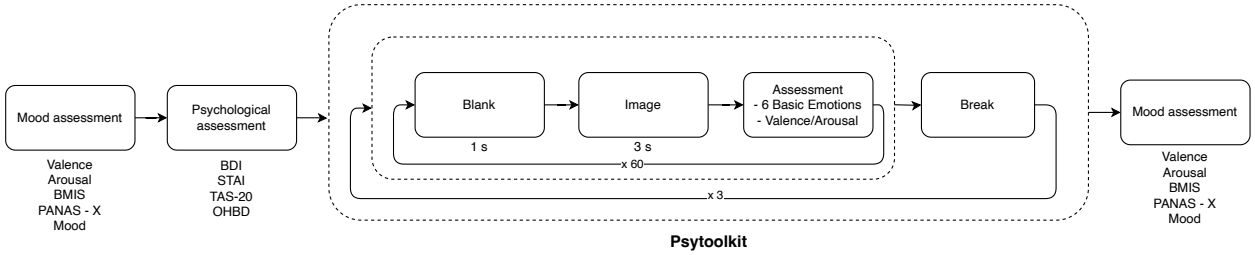


(a)

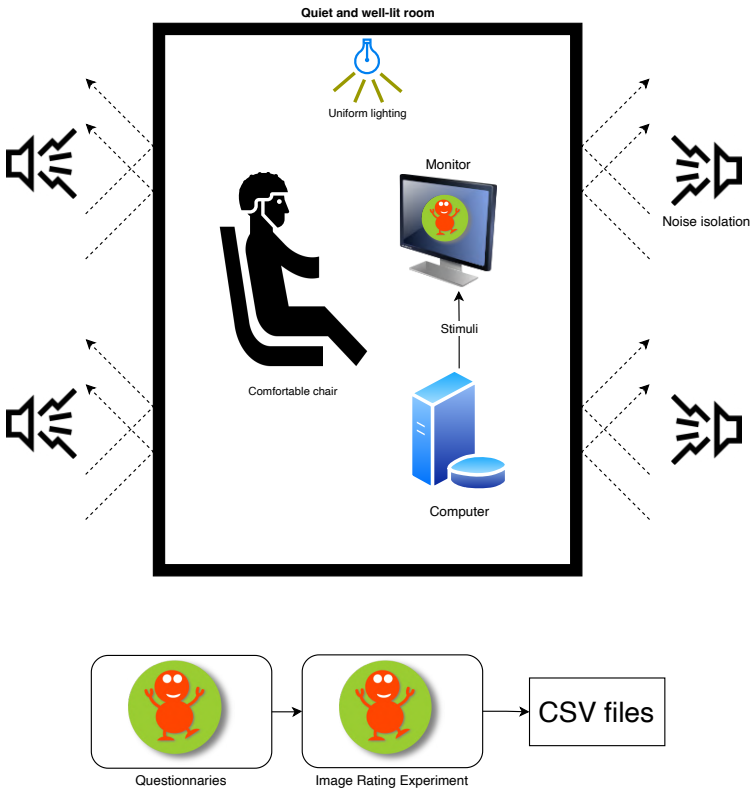


(b)

Figure 6.26: (a) The paradigm and (b) environment of the first part of the experiment. The questionnaires are explained in Appendix B.



(a)



(b)

Figure 6.27: (a) The paradigm and (b) environment of the second part of the experiment. The questionnaires are explained in Appendix B.

emotions: "happiness," "sadness," "fear," "surprise," "anger," and "disgust," on a 1 to 7 scale, following the same paradigm used to validate the NAPS-BE [378]. They were also asked to rate their level of valence and arousal on a 1 to 9 scale, consistent with the procedures described in the original studies on the NAPS-BE and SFIP. At the end of the online experimental paradigm, participants completed the following scales again to assess potential psychological variations due to completing the questionnaire:

- PANAS-X
- Arousal and Valence on a 1 to 9 Likert scale;
- BMIS;
- Mood on a -10 to 10 Likert scale.

Figure 6.27 shows the experimental paradigm of the online test and the environmental conditions required for participants.

6.4.4 Data preprocessing

For this experiment, 20 people were recruited, with an average age of 28 years, of whom 53% were women and 46% had a high school diploma as their highest qualification. For the questionnaires data, in the Appendix C the results of the questionnaries are shown. The preprocessing of the data was structured on the basis of recent scientific literature. The XDF files containing the PsychoPy label information and the time-series of the signal EEG were merged into a python object for event creation and filtered between 1 and 100 Hz. Subsequently, early stage PREProcessing (PREP) [142] was used for the removal of noisy channels, for their interpolation and for the application of Robust CAR, a procedure in which CAR is only applied on channels that are not excluded due to the excessive noise present. Finally, D.C. noise was removed in the frequencies 50 Hz and its multiples up to 250 Hz. In the next step, the ICs were extracted by ICA using the Extended Infomax [211] algorithm, and the artefact components were removed using the automatic ICs classification system IC-label [85]. Finally, the individual epochs of the "happiness", "neutral", and "sadness" events were extracted by including the baseline consisting of the black screen during the second preceding the image (-1000 ms), until the end of the stimulus preset three seconds after its appearance (+3000 ms). This pipeline was performed using MNE python library for EEG/MEG data analysis [255] for Python and with EEGlab [139] for Matlab [386]. The pipeline diagram is shown in Fig. 6.28. In this study, the baseline was not subtracted from the signal as an analysis of pre-stimulus events is crucial, with a specific focus on frequencies and areas related to the control of attention. A total of three identical images for each participant in the H and N categories were excluded, resulting in 3540 samples for analysis.

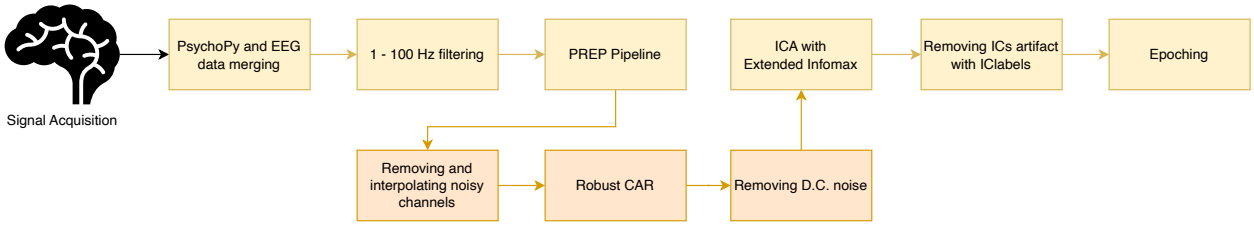


Figure 6.28: The preprocessing.

6.4.5 Feature extraction and preliminary analysis

For this thesis, only the EEG data collected in the first phase were used, and in future also the data from the second phase will be analyzed. Many factors were analyzed, from which it is possible to carry out a considerable amount of double and multiple comparisons. The selection parameters were the following:

- Time interval
- Frequencies (in the case of topographical analyses)
- Brain area
- Experimental conditions

Based on the literature illustrated, it is possible to state that in image viewing the first second of observation from the onset of the image is relevant, and in the studies analyzed baseline reduction is not performed. The initial components are the occipital ones with the N170 and subsequently the P200 and P300: the first moments are therefore relevant, and for this reason in the topographical analyses (Section 6.4.5.3) it was decided to analyze only the first 600 ms, since due to the nature of the calculation the temporal information is compressed to bring out the spatial one, while as regards the analyses that allow a view of the temporal dimension (Sections 6.4.5.1 and 6.4.5.2) it was decided to analyze from -200 ms before the onset of the image to +600 ms after its presentation. It is also known from the literature that negative images have a more intense response than positive and neutral images, and for this reason, for this preliminary study

Study Design	Acronym	Features	Statistical Test
happiness - sadness	HS	ERP ERSP Topographical ERSP	t-test
sadness - neutral	SN		t-test
happiness - neutral	HN		t-test
happiness - neutral - sadness	*HNS*		Analysis of Variance (ANOVA)

Table 6.15: Study designs. In *italic* the design used in this section.

Features	Data Domain	Analysis	Fixed Parameter
ERP	Areas: F_r, F_l, C_r, C_l, P_r, P_l	Time	Space
ERSP	Areas: F_r, F_l, C_r, C_l, P_r, P_l	Time-Frequency	Space
Topographical ERSP	Frequency bands: θ, α, β, $\downarrow \gamma$ Time intervals: 0 - 0.15 s, 0.15 - 0.3 s, 0.3 - 0.45 s	Space	Time-Frequency

Table 6.16: Analysis performed on the data.

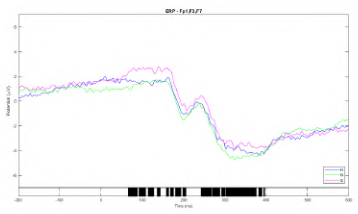
it was chosen to analyze the differences between negative images, "sadness" with positive ones, "happiness" and "neutral" (HNS) (Tab. 6.15).The analyses were carried out by analyzing both the ERP and the ERSP, both maintaining the temporal domain and synthesizing the spatial one, and subsequently the inverse using the topographical maps also illustrated in Chapter 3. In summary, the following sections refer to Tab. 6.16. Furthermore, it was chosen to analyze the brain areas by synthesizing the information of the channels without making a punctual analysis: in this way it was possible to reduce the noise by improving the SNR. The areas were briefly described in Chapter 2 and are as follows:

- (F_r) Right Frontal (Fp2, F4, F8)
- (F_l) Left Frontal (Fp1, F3, F7)
- (P_r) Right Parietal (P3, P7)
- (P_l) Left Parietal (P4, P8)
- (C_r) Right Central (C3, FC1, FC5)
- (C_l) Left Central (C4, FC2, FC6)

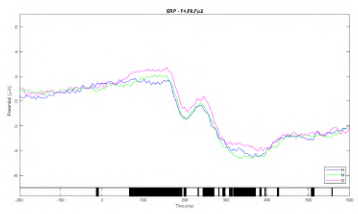
The signal will be analysed both in the time domain and in the frequency domain using topographic maps, as well as through time-frequency diagrams. The time domain analysis (ERP) will be performed by subdividing the data into the brain areas listed above, studying the signal between -0.2 and +0.6 s with respect to the appearance of the image, and the same will be done for the time-frequency diagrams (ERSP); while for the study in the frequency domain using topographic maps (Topo-ERSP) the following bands will be analyzed: theta band (4 - 8 Hz, θ), alpha (8 - 12 Hz, α), beta (12 - 30 Hz, β) and low gamma (30 - 60 Hz, $\downarrow \gamma$), illustrated in Chapter 3, in three different intervals: between 0 and 0.15 s, between 0.15 and 0.3 s and finally between 0.3 and 0.45 s. For all statistical analyzes, a significance level of $\alpha = 0.05$ was chosen as the threshold value.

6.4.5.1 ERP analysis

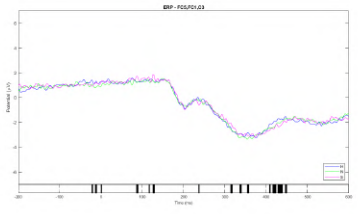
As already illustrated in the introduction, ERP are deflections of the EEG signal that are very small compared to the background rhythm. To bring out this information it is therefore necessary to make a point-wise average of many trials acquired under the same conditions: since the background rhythm is random, its average will tend to zero showing the constant information within the category of trials analyzed. The preliminary analysis of the ERP, extracted for the areas of interest in the time interval -200 and + 600 ms



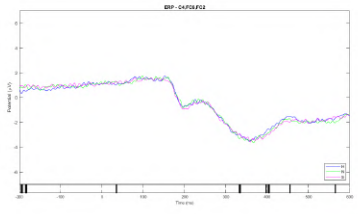
(a) F_l



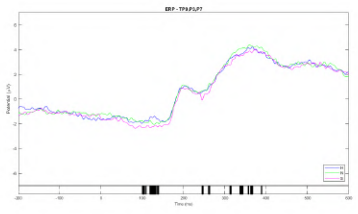
(b) F_r



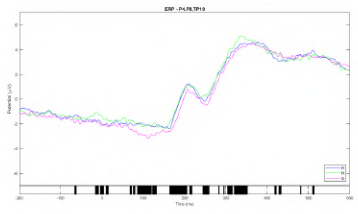
(c) C_l



(d) C_r



(e) P_l



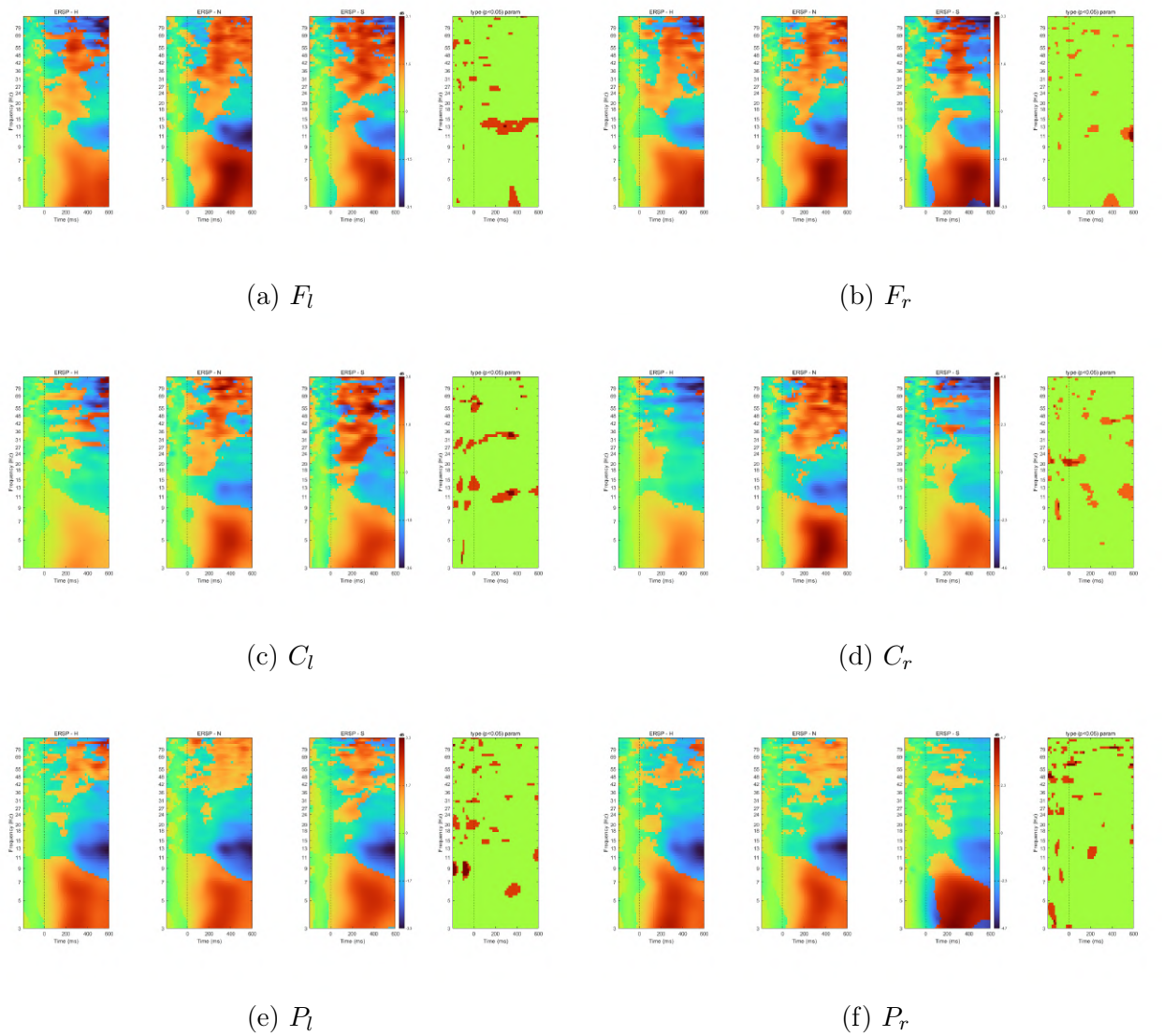
(f) P_r

Figure 6.29: Results for statistical test with $p - value < 0.05$ of ERP. The time interval is between -200 ms and +600 ms for HNS condition. (a) Frontal left, (b) Frontal right, (c) Central left, (d) Central right, (e) Parietal left, (f) Parietal right.

in the HNS condition is visible in Fig. 6.29. As can be seen from the statistic result, the significant values ($p - value < 0.05$) of the amplitude are all located after the onset of the image: this indicates that the statistically significant difference is given by the stimulus itself, and being the physical stimulus the same (both images appearing on the screen), it is possible to state that the difference in the amplitude of the ERP is given by the cognitive analysis of its content. Moreover, the P300 is clearly visible in the Fig. 6.29 (a) and (b), that corresponds to the (a) left and (b) right frontal lobe ERP, and this results are in line with literature. The "sadness" condition has in both case the most higher amplitude, and the amplitude in those peaks are statistically different from the other conditions. Regardless, additional post-hoc analyses are essential to more definitively ascertain which specific condition differs from the others.

6.4.5.2 ERSP analysis

The analysis of ERSP, computed using the Morlet Wavelet transformation, was performed for the regions of interest described previously for the HS condition in the time interval from -200 ms to +600 ms within



(a) F_l (b) F_r

(c) C_l (d) C_r

(e) P_l (f) P_r

Figure 6.30: Results for statistical test with $p - value < 0.05$ of ERSP. The time interval is between -200 ms and +600 ms for HNS condition, and the frequencies analysed goes from 1 to 100 Hz. (a) F left, (b) F right, (c) C left, (d) C right, (e) P left, (f) P right.

the 1 to 100 Hz frequency band, and a threshold of $p - value < 0.05$ was chosen for the statistical analysis. In Fig. 6.30 it is possible to observe the results of this preliminary analysis: it is clearly visible in Fig. 6.30 (a) and (b) that for the frontal lobe there is a marked difference in the theta and alpha band especially in the late components that indicate a cognitive processing of the image, but also in the area corresponding to 250-500 ms, where in Fig. 6.29 the deflection of the potential corresponding to the P300 is clearly visible.

6.4.5.3 Topographical Event-Related Spectral Perturbation (Topo-ERSP) analysis

The difference in the spectral power distribution on the scalp in theta, alpha, beta, and low gamma frequency bands was analysed using Morlet Wavelet decomposition, and tested with ANOVA and choosing a threshold value of $p - value < 0.05$. Moreover, three different time windows were analysed: 0 - 150 ms, 150 - 300 ms and 300 - 450 ms, shown in Fig. 6.31, 6.32 and 6.33, respectively. As can be seen, only occipital channels in alpha and beta frequencies bands, between 150 - 300 ms and 300 - 450 ms respectively, shown some differences. Although they are not statistically significative, marked frontal activation can be observed, particularly on

the left, as shown in previous sections with ERP and ERSP.

6.4.6 Training using ML and DL architecture

Preliminary analysis using the ERP are extremely useful to investigate the neuropsychological phenomenon present within the data, which, as illustrated previously, must be extracted through the averaging technique. However, the use of such neurocognitive information is not possible to guide a BCI, as it does not involve the use of numerous trials, but a single sample that must be classified. It is necessary to use classification techniques that can determine, on the basis of a single sample, trained on data previously provided to the classifier, whether the sample given as input is part of class H or S. The data used were pre-processed according to the pipeline illustrated above. For each technique, the following tests and countertests were performed.

Comparison 1: Images H vs N vs S The first test was to determine whether the AI system was able to distinguish between the two emotions, and according to the literature, the data were filtered with a bandpass filter from 1 to 60 Hz.

Comparison 2: Image VS Blank screen The first counter-evidence is to determine whether the system is classifying a visual stimulus, in this case the written image as shown in Fig. 6.26 (a) against the black screen. This is to determine whether the system actually classifies the emotion or simply a visual stimulus.

6.4.6.1 Machine Learning

First, the two different ML methods were used, both supervised and unsupervised: linear SVM, KNN, and K-means verifying the results based on the chance threshold, as defined in [145]. For these tests, the data were used by extracting the PSD with Welch's method in the theta, alpha, beta, and low gamma bands. It was observed that K-means does not cluster emotions in any of the cases, while it was observed that KNN and linear SVM are able to separate the samples in Comparison 2. In Tab. 6.17 the results are shown, and in Fig. 6.34 the confusion matrices are shown.

Model	Accuracy	Precision	Recall	F1-score
KNN	0.75	0.73	0.76	0.74
Linear SVM	0.84	0.87	0.80	0.83
K-means	0.49	0.49	0.97	0.65

Table 6.17: Results for Comparison 2

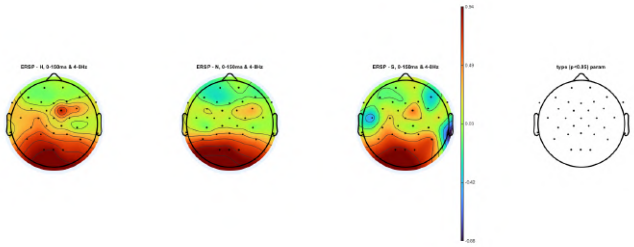
6.4.6.2 Deep Learning

For this training, the best architecture found in the previous study presented in Chapter 4 was used. It was observed that the network presented in Chapter 4 was able to correctly classify the samples only in the condition shown in Comparison 1, but not in Comparison 2. As an example, the results obtained using the EEGnet are reported: the results show that the EEGnet converges and the confusion matrix is shown in Fig. 6.35 (a) and the (b) loss curve on the validation set and (c) the training curve on the validation set are shown. In Tab. 6.18 the metrics for the same experiment are reported.

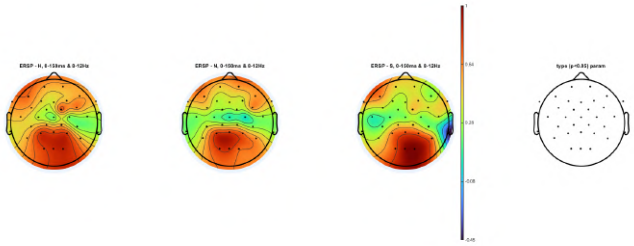
The results show that the classification was possible only in the Comparison 2 using some ML methods with frequency bands as features, and through the EEGnet using the signal in the time domain.

Based on these results, different explanations are possible, summarized below:

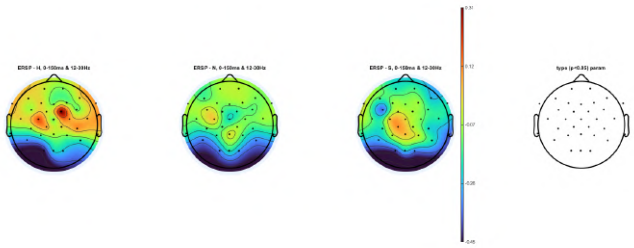
1. the experimental protocol is not adequate;



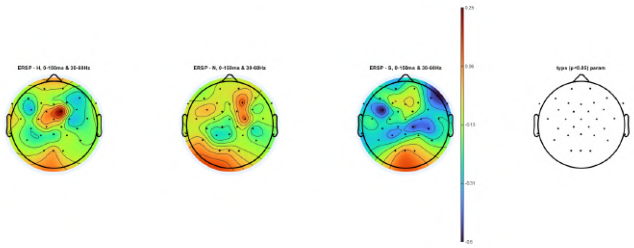
(a)



(b)

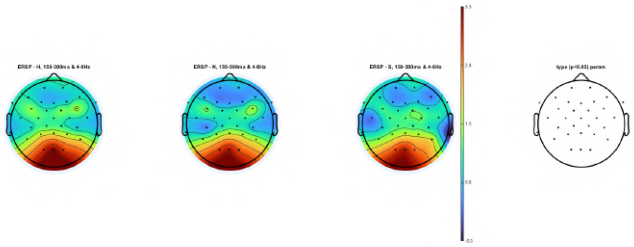


(c)

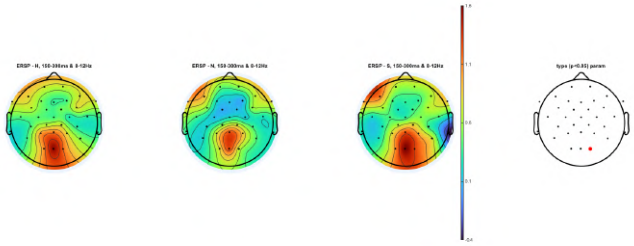


(c)

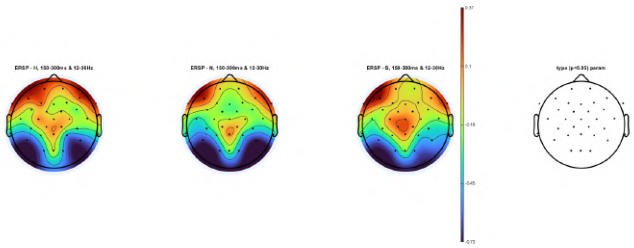
Figure 6.31: Topographical ERSP representation and statistical test between "happiness", "neutral" "sadness" (HNS) conditions in the following bands (a) theta, (b) alpha, (c) beta and (d) low gamma. The red electrodes indicates $p - value < 0.05$. The time interval is between 0 an 150 ms.



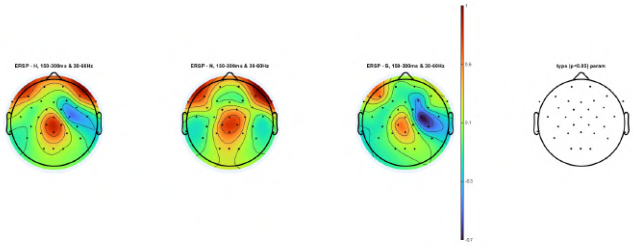
(a)



(b)

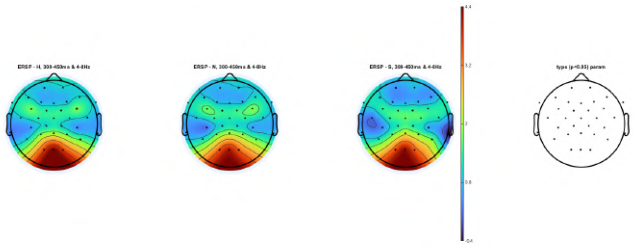


(c)

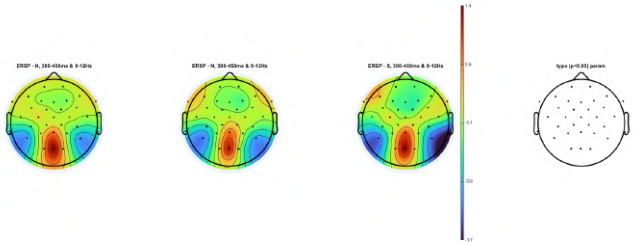


(c)

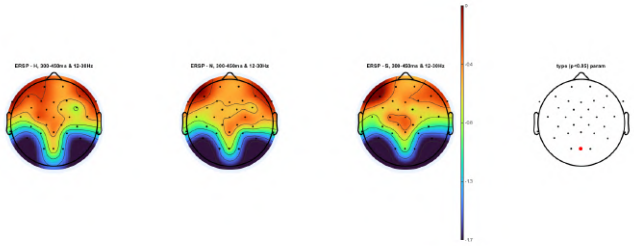
Figure 6.32: Topographical ERSP representation and statistical test between "happiness", "neutral" "sadness" (HNS) conditions in the following bands (a) theta, (b) alpha, (c) beta and (d) low gamma. The red electrodes indicates $p - value < 0.05$. The time interval is between 150 an 300 ms.



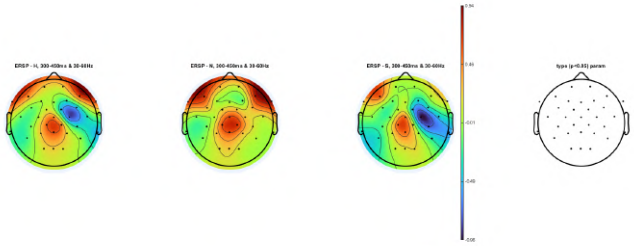
(a)



(b)

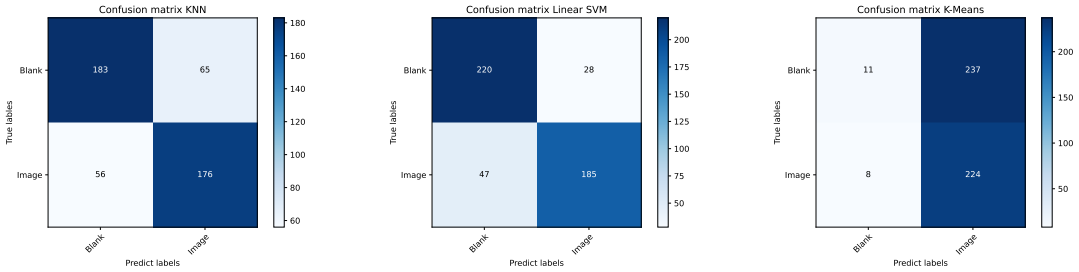


(c)



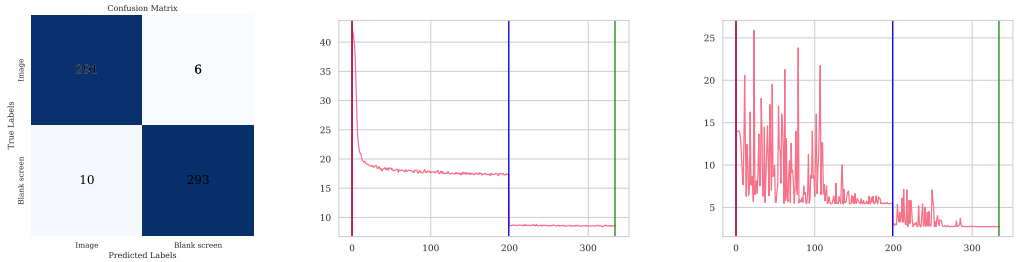
(c)

Figure 6.33: Topographical ERSP representation and statistical test between "happiness", "neutral" "sadness" (HNS) conditions in the following bands (a) theta, (b) alpha, (c) beta and (d) low gamma. The red electrodes indicates $p - value < 0.05$. The time interval is between 300 an 450 ms.



(a) (b) (c)

Figure 6.34: Confusion matrices results for Comparison 2 (Image VS Blank screen) for (a) KNN, (b) linear SVM and (c) K-means.



(a) (b) (c)

Figure 6.35: The (a) confusion matrix and (b) accuracy and (c) loss on validation set during the training with Image vs Blank Screen classes.

2. the sample size may be too small to train an AI system for classification;
3. the signal is too noisy;
4. the signal does not contain the desired information because emotions are not detectable via EEG.

Future developments should include the analysis of the psychometric data collected through the online and on-site questionnaires and check for possible relationships with brain responses. In addition, the same analyses will have to be carried out using the emotion, valence and arousal ratings given by the participants in the second phase of the experiment, as a new label of each image. Lastly, it will be necessary to analyse the hemispheric differences between the right and left hemispheres in the presentation of images, especially for the frontal lobe. Moreover, in future studies, the questionnaires and labels provided by the study participants will have to be analyzed as an alternative to the official labels of the two image datasets used. Finally, the classification tests will have to be repeated using non-brain signals (>60 Hz) to effectively verify the results obtained. In conclusion, these results are not in line with many of the studies published in the literature, but the refutation or confirmation of these results requires numerous tests that will have to be carried out in future studies. Future studies should conduct a deeper examination of ERP, ERSP, and Topo-ERSP in Comparison 2. The results are encouraging, if not excellent from a statistical point of view, and future investigations are necessary to better investigate the information contained in this experiment.

6.5 REMEMO Dataset

In this study, a novel EEG recording dataset is presented centered on self-evoked emotions. Twelve participants were recorded using a 64 channel EEG instrument while engaging in a slideshow. During the presentation, participants were asked to recall seven distinct situations from memory, each designed to elicit a specific emotion, and one neutral situation. The “neutral memory” is a novelty that is introduced to have

	Precision	Recall	F1-score	Support
Image	0.97	0.98	0.97	297
Blank screen	0.98	0.97	0.97	303
Accuracy			0.97	600
Macro avg	0.97	0.97	0.97	600
Weighted avg	0.97	0.97	0.97	600

Table 6.18: Metrics for Image vs Blank Screen classes experiment with EEGNet

a baseline for each emotion evoked. Before recording EEG, the participants completed a comprehensive psychological battery questionnaire. Subsequently, at the end of the recording sessions, the participants were administered SAM [314] to further evaluate their emotional states and another short questionnaire to evaluate the success of each evoked emotion. This methodology provides a unique contribution to the field, emphasizing the importance of a neutral baseline for the accurate interpretation of emotional responses recorded through EEG.

6.5.1 The experimental paradigm

The objective of REMEMO is to create a dataset in which subjects self-evoke emotions in the almost total absence of external stimuli. It was designed to contribute to emotion research by filling in the biases of previously published datasets listed in Tab. 6.2. In this experimental protocol, each participant feels seven self-stimulated emotions five times, based on their memory. Before all stimuli were presented, the fixation-cross was used to maintain the subject's attention and to record a baseline. The novelty introduced in this section is the recording of "neutral memory recall" after each single emotion: considering that the emotions evoked by the participants are based on their personal memory recall, the aim is to record a signal with the same pattern for the recalling task but without any affective valence. Each participant was trained to recall one memory for each type of emotion and to recall one neutral memory. At the end of the whole task, a SAM was filled in for each type of emotion asking the participant to select the values based on the average of the five repetitions of that emotion. Moreover, after SAM, the subject had to indicate whether or not he/she had success in this evoking. Before the experiment, each participant filled out psychometric questionnaires, fully explained in the Appendix B:

- Question on current mood (before experiment);
- Pittsburgh Sleep Quality Index (PSQI) [387];
- State-Trait Anxiety Inventory (STAI) [388];
- Beck Depression Inventory (BDI) [389];
- Toronto Alexithymia Scale (TAS) [390];
- Open Hemispheric Brain Dominance Scale [391];
- Millon Clinical Multiaxial Inventory (MCMI-III) [392].

The multiple repetition of the same task, the multiple session, and multiple subjects allow the researchers to investigate the stability of the emotional pattern on the subject and across subjects. The role of "fixation-cross" is to capture the subject's attention on the screen and to record the baseline. In addition, it also contains information on valence, arousal, and dominance related to the mood the subject has at the time of recording. In addition, the fixation-cross is common to many other datasets and allows for easier comparison. Our hypothesis about the translation of the Russell Plane due to mood is depicted in Fig. 6.37: the red arrow represents the influence of mood on the emotions recorded. In this hypothesis, emotions are translated by an unknown value on the axes corresponding to the mood of the subject at the recording time. The role

of the "neutral memory recall task" is to capture the neural pattern of "memory recall" and the "reading process" without any affective influence. The identical structure of the slide with emotional words allows us to compare the information between neutral and affective tasks. The hypothesis about the interaction between memory recall and emotion evoking is that the interaction is not the same in all frequencies and channels, as shown in Fig. 6.36. The graph represent a possible distribution of three distinct emotions within a categorical system. Color coding represents emotional intensity (Valence) on a scale from 0 to 1. The horizontal axis defines the channel space, ranging from 1 to c, while the vertical axis represents the frequency space, from 1 to f. It is hypothesized that emotions can be mapped within a three-dimensional space defined by these parameters. A further hypothesis is that, once encoded, the identification of a specific emotion can be achieved through a classifier trained to recognize the distribution of emotional parameters within this space. This classifier would be able to selectively analyze the channel-frequency combinations relevant to the target emotion, identifying its specific trend within the three-dimensional space. The REMEMO uses the fixation cross as an interval between different emotions that makes the REMEMO comparable with most other datasets, and neutral tasks as baseline. The signal recorded during the fixation-cross also contains physiological information about the subject's mood that could shift the show Russel Circumplex Plane, as shown in Fig. 6.37. Furthermore, the neutral task is used as a baseline since it is comparable in duration, type, and elements to the emotional task (20 seconds of duration, white writing on a black background, memory task, fixation cross before the task) but does not contain the emotional component. In this way, it is possible to subtract the information from the neutral task to uncover the relevant features related to the emotion. Furthermore, this dataset stores, for each person, psychological information on depression, anxiety, sleep quality, and personality evaluation. Moreover, information on the dominance of the brain hemisphere was acquired and it could be interesting to consider the hypothesis of lateralization of emotions [393]. The EEG data were acquired using a 64-electrode high-density instrument with a high sampling rate (1000 Hz) in a controlled environment, as shown in Fig. 6.39. The single source stimulation (slides) used as a trigger has low intrusiveness, as it is composed only of a white fixed text on black background without movements, variations in lights and colors, and a total absence of auditory stimulation. Furthermore, there is no interaction between the subject and the acquisition system, since the psychometric evaluation is performed before the task, the global evaluation of the task is carried out only at the end of the task, and the stimuli have the same duration to improve the comparability of statistical analysis. In this database there are subjects without neurodegenerative diseases, brain injuries, or subjects who use drugs. Furthermore, no minimum level of education is required to perform the task, and this is a novelty introduced with this dataset, making it more generalizable. This dataset has at this moment a low number of subjects, which will increase in the following months, transforming it into a multicentre dataset due to the sharing of the materials and methods used, to build a representative sample of the population. Finally, the multiple recordings of each subject make this section a prospective study that aims to extract individual patterns and analyze the individual's emotional development across their lifespan. The experiments were carried out in an isolated room in the Department of Biotechnological and Applied Clinical Sciences. In Fig. 6.39 is shown the environment composed of a computer that controls the EEG instrument and a screen for the presentation of stimuli. The participant was seated in a comfortable chair in front of a 20" high-resolution monitor. Each subject was trained the week before to prepare an emotional memory to better perform the task. For each subject, two sessions are recorded with at least one week between them. During EEG setup, the questionnaires are filled out on the RedCap[9] online platform. The experimental protocol is shown in Fig. 6.38. The participant is exposed to the vision of a slide show in which the emotion or neural memory to be evoked appears. The emotions used are shown in the Tab. 6.19. Each fixation cross was shown for 5 seconds, followed by the word (emotion or neutral) shown for 20 seconds. Finally, after the whole session, a self-assessment with SAM [314] is presented for Valence, Arousal, and Dominance, and the participant must complete a questionnaire SAM for each type of emotion. The order is the same for all people, but is different

[9] `https://www.project-redcap.org/`

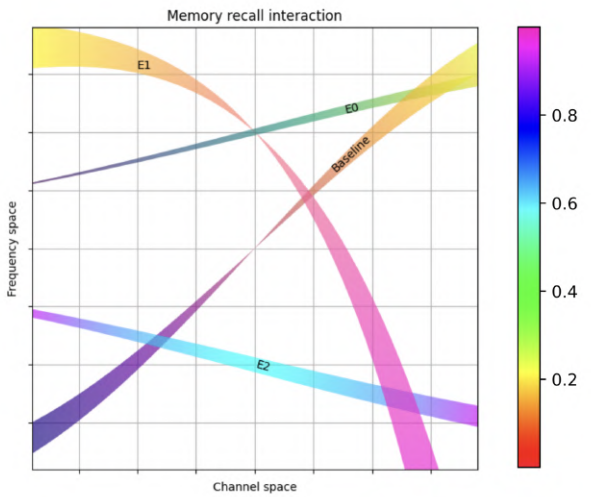


Figure 6.36: An example of the interaction between memory process (*baseline*) and the emotions on frequency-channel plane. Each emotion has a specific areas in which, with more or less intensity coded by color, it act at specific place and frequency. The crossing between the baseline and the emotions represent the most difficult point of recognition due to the interference. The colours in the graph have been made more transparent than in the color bar to highlight any overlaps between the different ribbons.

between the two sessions. All participants were Italian native speakers. The emotions listed on Tab 6.19 were chosen to cover all Russell Circumplex Plane [302] shown in Fig. 6.37 and these emotions are also the basic emotions studied by Ekman [394].

Valenza \Arousal	**High Arousal**	**Low Arousal**
High Valence	Happiness; Surprise	Relax
Low Valence	Anger; Fear	Disgust; Sadness
Neutral	Memory	

Table 6.19: Emotions categorized by Valence and Arousal.

Contrary to the datasets previously published in the literature, which are extremely useful for studying differences between emotions with different methods, the first purpose of the REMEMO is basic scientific research. To achieve this goal, a new experimental protocol was introduced in which each participant remembers and attempts to feel each emotion more than once in a single recording session, and each session is recorded twice. To record the most cleaned data, two stimuli for controlling the "cognitive noise" were introduced:

- *fixation-cross*;
- *neutral memory recall task.*

6.5.2 Preprocessing

The preprocessing of REMembrance of EMOtions (REMEMO) has a slightly different structure than the studies presented previously. Unlike MI/ME and sleep, the EEG emotions data have a greater component of subjectivity compared to movement, and even in the literature there is no broad agreement [10] on which channels or frequencies are relevant for the analysis of emotions through EEG.

Data were acquired from 17 subjects, and each subject was subjected once to the experimental protocol shown in Fig. 6.38, recording a total of 7 self-evoked emotions 5 times each, and 35 neutral memory recalls, one for each emotion. For this preliminary analysis, only the data of the emotions "happiness" (H) and

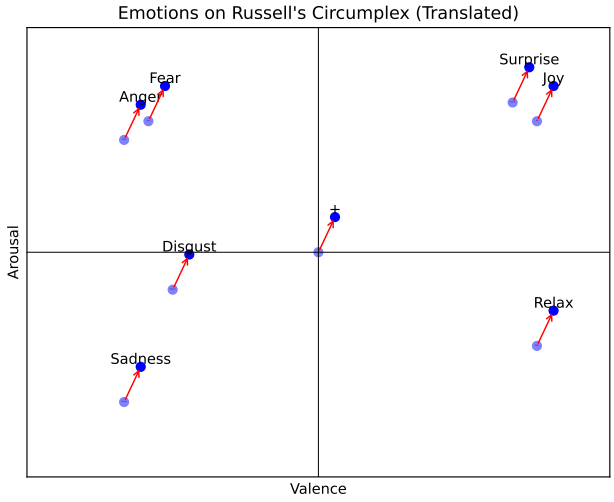


Figure 6.37: An example of the spatial distribution of the emotion used in REMEMO on Russell Circumplex Plane. The transparent dots represent the theoretical place of the emotions, while the translated dots represent the emotions recorded that are translated by ϵ depending on the mood at the time of recording. The fixation cross, with no emotional relevance, is translated only by the value of the mood. The red arrow represents the translation vector, which is the mood of the participant at the recording time.

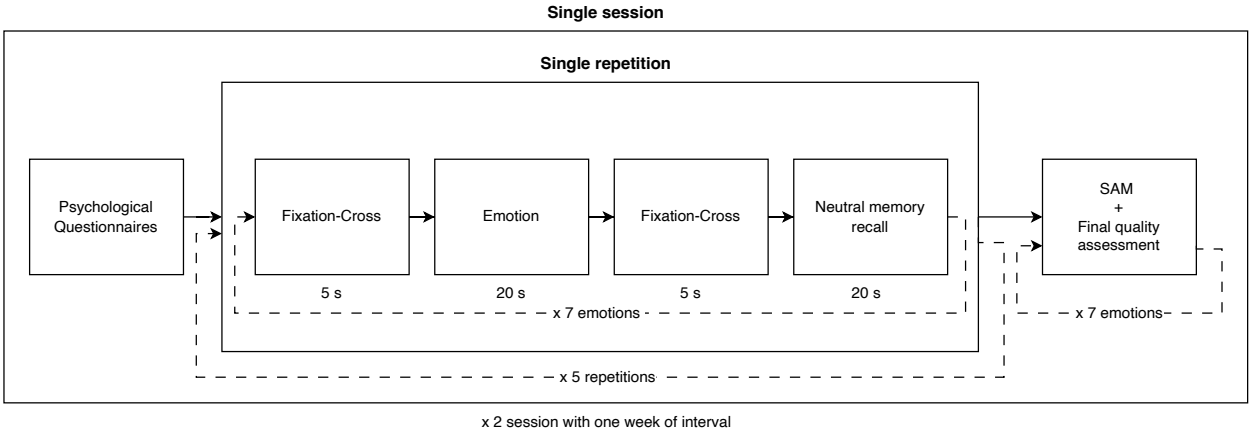


Figure 6.38: Experimental Protocol for the REMEMO dataset. The questionnaires are explained in Appendix B.

"sadness" (S) were analyzed and comparisons were made between the two emotions using the t-test (t-test) and choosing $\alpha = 0.05$.

Data were initially filtered between 1 and 100 Hz, and subsequently the PREP preprocessing pipeline was applied to remove DC noise, noisy channels, and their interpolation and to ensure robust CAR. Subsequently, the signals of the 64 electrodes were decomposed via ICA using the Extended Infomax algorithm [211], and the components classified as artifacts by IClabel [85] with a probability greater than 90 % were removed. The entire preprocessing pipeline used is shown in Fig. 6.40.

6.5.3 Descriptive analysis on ERP

The signal was divided into epochs and the following features were extracted:

- ERP and Topo-ERP for the left temporal lobe (text analysis);
- Topographical Event-Related Spectral Perturbation (Topo-ERSP) to verify any statistical differences between emotions in individual electrodes.

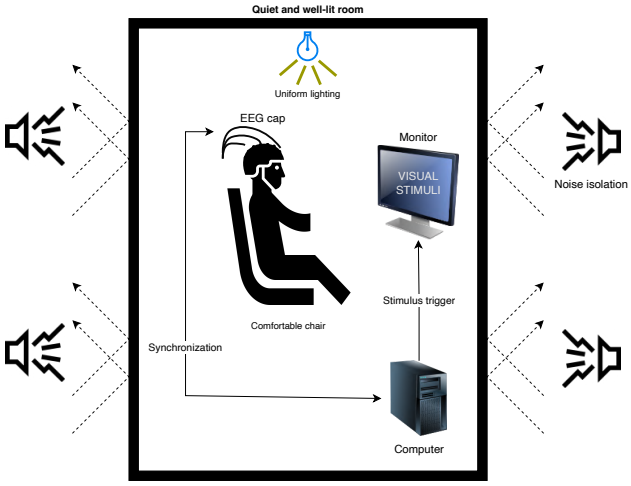


Figure 6.39: Environment isolated from noise and external stimuli with a uniformly lighting.

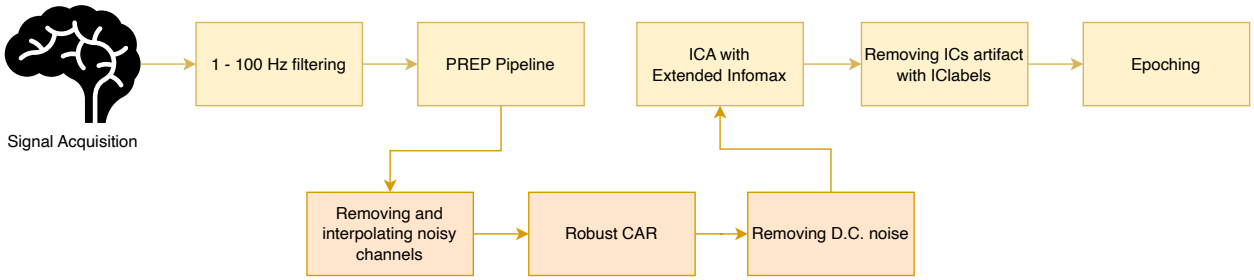


Figure 6.40: The pipeline used for preprocessing.

Analysis using ERP is a well-known technique in the literature and allows, through averaging, to extract components of the EEG signal that are repeated in all trials, eliminating the background rhythm that tends to zero when averaged. In the dataset presented here, as shown in the Tab. 6.2, REMEMO does not use stimuli that need to be analyzed but only *cues* that indicate which memory, known only to the participant, they must recall from memory. For this reason, the hypothesis of this preliminary analysis is that ERP does not show any type of cognitive processing outside of the processing of the written word (N400) in the central-parietal lobe [395]. On the contrary, the objective is to verify if within the same class of cross-subjects trials or within the repetitions of the same subject, there are invariants that can instead emerge through Topo-ERSP.

From the analysis of the literature, it emerges that for tasks of recalling emotional memories, in particular images, [396] investigates how the brain behaves during learning and memory recall. To do this, the authors used multimedia content and high-resolution EEG recordings and recorded the memory recall process. In particular, they observed that successful memory recall was associated with specific brainwave patterns: theta waves (4 - 8 Hz) in the frontal and parietal regions and gamma waves (31 - 48 Hz) in the left posterior temporal and frontal regions. For this reason, preliminary analyzes were carried out taking into account the results of this study, which shows points in common with the experimental protocol of this section.

6.5.3.1 ERP e Topo-ERP on centro-parietal lobe

The analysis of ERP and Topo-ERP for the identification of N400, carried out preliminarily on all the stimuli of "happiness" (H) and "sadness" ($2_{emotions} \times 5_{repetition} \times 17_{subjects}$) stimuli for a total of 170 stimuli, 85 for each emotion. The information of all the trials together was extracted relative to the central-parietal channels (CP5, CP1, CP6, CP2, CP3, CPz, CP4), and the result is shown in Fig. 6.41 (a) and (b), but no component is visible, probably due to temporal subjectivity in memory recall.



(a) (b)

Figure 6.41: (a) ERP for centro-parietal, between -200 ms and 800 ms, (b) Topo-ERP between 0 ms and 800 ms.

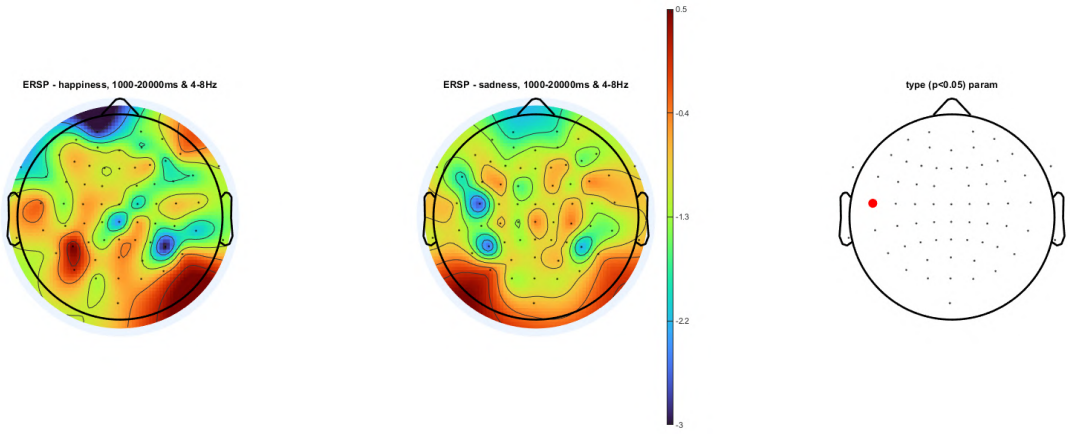


Figure 6.42: Topo-ERSP for whole period excluding the first second avoiding the influence of the reading, for frequency band between 4 - 8 Hz. The threshold for significance is $p - value < 0.05$.

6.5.3.2 Topo-ERSP

Spectral analysis was performed using Wavelets, which, as introduced in Chapter 3, are particularly useful for analyzing EEG signals because they allow highlighting signal features with different scales for each frequency. For the analysis of this signal, the bands shown in the literature between 4 and 48 Hz [396] were analyzed, and the results are shown excluding from the calculation the period before the onset of the image, separating the 4 - 8 Hz band (Fig. 6.42) from the 31 - 48 Hz band (Fig. 6.43), as defined in [396]. The $p - value < 0.05$ was chosen as the threshold value for statistical significance.

6.5.4 Training using ML and DL architecture

Preliminary analyzes using ERP are extremely useful for investigating the neuropsychological phenomenon present within the EEG data, which as previously illustrated must be extracted through the averaging technique. However, the use of such neurocognitive information is not possible to guide a BCI, since it does not involve the use of numerous trials, but a single sample that must be classified. It is necessary to use

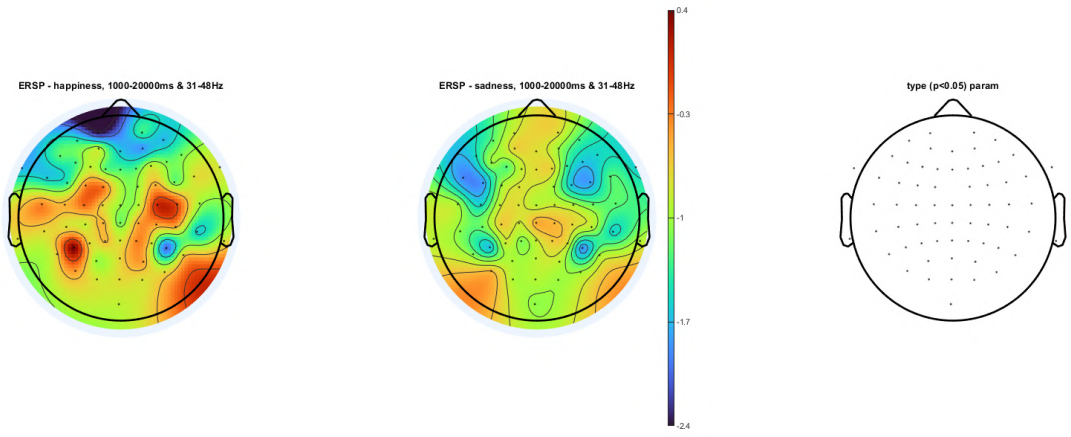


Figure 6.43: Topo-ERSP for whole period excluding the first second avoiding the influence of the reading, for frequency band between 31 - 48 Hz. The threshold for significance is $p - value < 0.05$.

classification techniques that can determine, based on a single sample, trained on data previously provided to the classifier, whether the given input sample is part of class H or S. The data used were preprocessed according to the pipeline illustrated previously. For each technique, the following tests and countertests were carried out.

H VS S emotion stimuli The first is whether the system was able to distinguish between the two emotions, and based on the literature, the data were filtered with a bandpass filter from 1 to 60 Hz.

Emotions stimuli (H+S) VS Non-Stimulus The first counter-evidence is to determine whether the system is classifying a visual stimulus belonging to class H or S, in this case the written text as shown in Fig. 6.38 with respect to the fixation cross. This is used to understand whether the system is actually classifying the emotion or simply a visual stimulus.

6.5.4.1 Machine Learning

Initially, three different ML methods were used, both supervised and unsupervised: SVM, KNN and K-means but no system in the trials was able to pass the chance threshold, as defined in [145]. For these tests, data were used by extracting PSD using Welch's method in the theta, alpha, beta and low gamma bands.

6.5.4.2 Deep Learning

For this training, the best architecture found in previous studies presented in Chapter 4 was used.In this preliminary analysis, the period between 1 and 20 seconds and all 64 channels was used. In none of the tests, the network was able to classify the signals correctly.

The REMEMO dataset, unlike previously published datasets that mainly rely on external emotional stimuli, focuses on internally generated emotions triggered by the recall of personal memories. A similar idea was explored with the IMAGINED EMOTIONS dataset [358], but it contained only one recording per emotion for each participant, making it difficult to study the intra-individual emotional pattern. This

approach, on the other hand, offers a unique opportunity to study the neural correlates of emotions in a more ecologically valid context, minimizing the influence of external sensory processing and revealing cleaner and less biased emotional responses. The main innovation of REMEMO is the introduction of a "neutral memory recall" condition, as occurs during the study of "resting-state" in the study of MI/ME. This serves as a crucial baseline, allowing one to isolate brain activity specifically related to emotional experience from general cognitive processes associated with memory retrieval that could be very subjective. This methodological advance addresses a significant limitation of previous studies, where differentiating emotional responses from cognitive load has been a challenge. The experimental design, characterized by multiple repetitions of each emotion within and across sessions, allows investigation of the stability of the emotional response both within and across individuals. This longitudinal approach allows us to analyze the dynamics of emotional processing over time and to identify individual differences in emotional reactivity and regulation, as well as the stability of a possible emotional pattern within the same individual. The absence of an N400 component in the centro-parietal ERP analysis associated with the semantic processing of writing suggests that the verbal stimulus presented during the task may not have affected the EEG data. The analysis using Topo-ERP showed brain activation patterns in different brain regions that could be associated with emotional processing. Time-frequency analysis using Topo-ERSP, focusing in particular on theta (4-8 Hz) and gamma (31-48 Hz) bands, revealed significant differences in brain activity between the conditions "happiness" and "sadness". These findings are consistent with previous studies linking these frequency bands to memory recall and emotional processing [396]. The negative result in the classification of different emotions using ML and DL techniques, both of different emotions ("happiness" vs. "sadness") could be attributed to several factors: the relatively small sample size (17 subjects) may have limited the statistical power of the analyses; the intrinsic complexity and variability of emotional experiences; the choice of features and model architectures may not have been optimal for capturing the subtle and complex neural signatures of self-generated emotions. This shows how, although the results presented by the studies published from the data analysis EEG acquired with external stimuli show very high classification metrics, the detection of self-induced emotions does not present the same results.

6.6 Discussion

This chapter has presented a series of experiments on EEG signals of emotions, both on public datasets and on collected data. Specifically, experiments were conducted both with visual stimuli and through data collected via a paradigm of self-induction of emotional states. Through the analysis of public datasets and the creation of two new datasets, innovative AI architectures, FC analysis methods, and emotion elicitation protocols were explored to reduce biases present in the data. The literature analysis highlighted the intrinsic problems in emotion recognition via EEG, including inter-individual variability of EEG signals, the complexity of human emotions, the quality of signals often contaminated by noise and artifacts, the ecological validity of experimental paradigms, the absence of a commonly shared theory supported by empirical evidence and certain neurobiological findings. Despite the advanced data analysis methods using AI techniques shown in the published literature and despite the promising results published in terms of performaces, their generalizability remains limited. In the literature on emotion recognition via EEG data, a notable increase is observed, with ever-improving results in terms of classification performance of two or more emotions. In [397], the authors classified emotions into three and four classes using SEED, presenting a CNN-Bi-LSTM architecture, with an accuracy of 99.55% and 99.79%, respectively. In [398], the authors used a Bi-Branch Vision Transformer network for signal classification using the SEED dataset, achieving an accuracy of 66.77% on four classes. In [399], a CNN-based architecture was presented that achieved performances of 96.43% and 96.30% in arousal and valence classification tasks using DEAP. The analysis of public EEG emotion datasets revealed the presence of biases related to elicitation protocols, the limited number of sensors, equipment stability, and the lack of an adequate baseline. These factors contribute to the difficulty of developing robust

and reliable models for emotion recognition. Furthermore, the lack of a theory from which stimulation protocols can be derived appears to be key, making it impossible to standardize them for a comparison of results. The exploration of innovative architectures, such as the use of ConvLSTM architectures for the classification of the "relax" state using the IMAGED EMOTION dataset [358], has opened new perspectives in the analysis of emotional EEG signals and introduced the concept of a reference emotion, as close as possible to a "neutral" emotional state, as occurs in MI/ME with the "resting-state". However, the preliminary results obtained with the 4D LSTM network showed considerable variability in prediction accuracy among different subjects, confirming the subjective nature of emotional patterns and the difficulty of generalizing results at the inter-individual level. The study of FC through PLI and its graphical representation provided interesting insights into the differences in the activation of brain regions and frequency bands associated with different emotions. The qualitative analysis of the graphs showed distinct connectivity patterns for emotions such as "anger", "happiness", "sadness", and "relax" suggesting the existence of specific emotional *fingerprints* for each individual. Although the graph-based approach represents a promising method for the analysis of brain connectivity, the current limitation of the number of subjects analyzed does not allow definitive conclusions to be drawn. Furthermore, this method of analysis will be used to verify whether CV architectures are able to extract patterns through this new visualization of FC. The two experimental protocols developed for data acquisition through visual stimuli and through self-induced memory (REMEMO) introduce innovative elements for emotion research. The REMEMO dataset focuses on the self-induction of emotions through the recall of personal memories, reducing biases related to external stimuli and offering a more natural context for emotion analysis. The introduction of a "neutral memory" as a baseline represents a novel element that could improve the accuracy of the interpretation of emotional EEG signals. The EEG data dataset created through visual stimulation, on the other hand, is based on the presentation of images with validated emotional valence (from NAPS-BE and SFIP) and introduces the collection of subjective evaluations of emotions by the participants, allowing both inter- and intra-subjective analysis. The preliminary analyses conducted on the acquired data provided mixed results. On the one hand, on the data collected with the visual stimulation protocol, the analysis of ERP and ERSP confirmed some patterns of brain activation consistent with the literature, such as the activation of frontal areas in response to emotional stimuli and differences in theta and alpha frequency bands. On the other hand, the application of ML and DL techniques for emotion classification produced results below expectations, in particular for the REMEMO dataset. These results suggest several possible interpretations. It is possible that the results are due to an inadequacy of the experimental protocol: despite efforts to reduce biases, the protocols used may not be optimal for the elicitation of self-induced emotions (REMEMO) or for the classification of emotions evoked by visual stimuli. Another limitation is the limited sample size, as the number of participants may be insufficient to effectively train artificial intelligence models, especially considering inter-individual variability. Furthermore, the quality of EEG signals may be influenced by artifacts or an unfavorable SNR, limiting the effectiveness of classification techniques. Finally, as already shown in other sections, it is possible that emotional patterns, especially self-induced ones, are not detectable via EEG signal since it is possible that cortical activation is too complex or variable to be captured by current models. In future studies, classification tests using non-brain signals ($>$100 Hz) will need to be repeated to effectively verify the results obtained to confirm that the signals obtained by filtering $<$60 Hz are actually the result of cognitive activity. Moreover, future research should focus on the analysis of individual differences, both in terms of brain activation patterns and psychometric responses, to better understand the subjective nature of emotions. Furthermore, the use of multimodal data such as EEG, fNIRS, and fMRI could provide a more complete and detailed view of the neural correlates of emotions. Finally, the development of emotional BCI systems, both active and passive, represents an ambitious but highly relevant application goal. The ability to recognize and interpret emotions in real time could open new perspectives in various fields, from clinical to affective robotics, from education to entertainment. The results therefore suggest that in the field of AC, and in particular of emotion recognition, it is appropriate to take into account the results of the counter-tests shown in the last sections to confirm that the classification is actually that of the EEG signal resulting from the emotional state.

Part III

Conclusions

Chapter 7

Conclusive Hints and future developments

This thesis aimed to present effective AI-based models to decode EEG signals related to complex brain functions. The final goal was to extract and decode emotions by improving neuroinformatics tools and BCI models. To achieve this goal, two fundamental areas of neuroinformatics have been explored using EEG data and DL methods: ME and the sleep domain. These two domains are well known in the literature in terms of neural processes and EEG signal patterns. In contrast, the emotion domain remains less understood regarding neural circuits and psychological models. Moreover, these seemingly distinct research fields share a common thread: the goal of developing lightweight, bias-free models that are easily applicable in the BCI domain. Special attention was given to states of nonaction (i.e., *resting state* for movement, *relax* for emotions, and SWS for sleep as a "bad" stage for awakening, causing sleep inertia), to promote the ecological and sustainable application of these models in real-world contexts.

The study of movement highlighted that ME still receives less attention compared to imagined movement. Furthermore, it was found that the state of non-movement, commonly referred to as *rest*, has been underexplored in existing research. The results of this work demonstrated that the integration of kinematic data can significantly enhance the accuracy of inferential models, opening new perspectives for practical and theoretical applications.

The study of sleep states emphasised the potential of DL models to predict transitions between sleep states using temporal series of EEG signals. Through a pilot study, it was shown that such models can contribute to optimising rest periods, offering significant implications for health and well-being.

A significant contribution of this work has been the analysis of emotions, which revealed that this field is among the most complex in neuroinformatics, primarily due to the intrinsically ambiguous and theoretically undefined nature of emotional data. This presents both a challenge and an opportunity for advancements in the neuropsychological domain and *Affective Computing (AC)*, stimulating further technological and theoretical developments. In addition, it encourages the improvement of experimental protocols for data acquisition, considering the physics of the devices used and the neurophysiology of the CNS.

Another important contribution was the creation of three new datasets, which will be made available to the scientific community to support the development of innovative models, validate existing methodologies, and expand knowledge in the explored domains.

Furthermore, it has been shown that DL can classify EEG signals with performance that can be improved by preprocessing and architectural optimization. However, the domain of emotions appears to be much more complex than the others analyzed in this thesis. Several published works show very high emotion classification performance in public datasets. However, they are affected by biases that could result in the classification of something different from emotions, such as the content of the stimulus cognitively processed by the subject. The brain regions associated with emotions are mainly subcortical, where EEG has very

limited sensitivity: EEG primarily records the electrical fields generated by layers III, V, and IV of the cortical system.
The differences observed in the results obtained with the DL methods in different datasets could be attributed to the type of stimulus used. For example, visual stimuli produce distinct patterns, as shown by analyses of ERP, ERSP, and their topographic representations, whereas nonvisual stimuli elicit comparable emotions but do not yield similar results with the same DL methods. This discrepancy could also be due to the sample size but is certainly influenced by the stimulation protocol, which plays a significant role in model performance.
It is crucial to underline several limitations. Despite promising results in controlled laboratory settings, it is essential to conduct research in ecological environments with high noise in order to develop systems that can be used in real-world scenarios. Furthermore, a distinction emerges between controlled task acquisition protocols and the use of systems in the real world. In order to generalise both the results and the users of the systems, it would be advisable to enlarge the sample of subjects, diversifying it according to age, gender and hand lateralisation. Lastly, it would be useful to test these models on other datasets, such as those described in Chapter 6, both during the training phase to improve performance and generalisability, and to test them with protocols, tools and subjects that differ widely in terms of culture, location and ethnicity.
These findings suggest that unlike other cognitive activities, such as sleep or movements, the study of emotions requires careful consideration of the stimulus and protocol before training AI models. This is because the data could reflect the cognitive processing of the stimulus rather than the currently elicited emotion. In addition, stimuli with unpleasant emotional content tend to increase attentional thresholds, which is reflected in certain EEG frequency bands. This effect, however, cannot be attributed to the emotion itself but to its cognitive processing and the increased attention to the stimulus. Therefore, it is necessary to continue refining experimental protocols for EEG data acquisition and to apply counter-tests to ensure that the classified states are emotional rather than cognitive. This fundamental step toward the creation of an ideal BCI will be the subject of future developments in light of the results reported here.
In conclusion, this thesis demonstrated how neuroinformatics can be a fundamental bridge between neuroscience and computer science, significantly contributing to the study of cognitive dynamics and paving the way for future developments that may improve the human conditions. Interdisciplinarity thus emerges as a crucial resource for advancing knowledge and applications in neuroscience.

Part IV

Bibliography

Bibliography

[1] M. Abrams and J. D. Van Horn, "Bridging the gap: How neuroinformatics is preparing the next generation of neuroscience researchers," 2024.

[2] G. A. Ascoli, D. N. Kennedy, and E. De Schutter, "Farewell, neuroinformatics!," 2021.

[3] R. Picard, "W.,(1997). affective computing," 1997.

[4] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. Siegelbaum, A. J. Hudspeth, S. Mack, *et al.*, *Principles of neural science*, vol. 4. McGraw-hill New York, 2000.

[5] D. Purves, G. J. Augustine, D. Fitzpatrick, W. Hall, A.-S. LaMantia, and L. White, *Neurosciences*. De Boeck Supérieur, 2019.

[6] N. R. Carlson, *Foundations of physiological psychology*. Pearson Education New Zealand, 2005.

[7] C. Bonfiglioli, U. Castiello, *et al.*, *Metodi di indagine in neuroscienze cognitive*. Piccin, 2005.

[8] U. Castiello, *Tecniche sperimentali di ricerca in psicologia*. Piccin-Nuova Libraria, 1995.

[9] G. Denes, L. Pizzamiglio, C. Guariglia, S. Cappa, D. Grossi, C. Luzzatti, *et al.*, *Manuale di Neuropsicologia. Normalità e patologia dei processi cognitivi (Terza edizione)*. Zanichelli, 2019.

[10] E. E. Smith and S. M. Kosslyn, "Cognitive psychology: Mind and brain," *(No Title)*, 2007.

[11] F. Zeidan, K. T. Martucci, R. A. Kraft, J. G. McHaffie, and R. C. Coghill, "Neural correlates of mindfulness meditation-related anxiety relief," *Social cognitive and affective neuroscience*, vol. 9, no. 6, pp. 751–759, 2014.

[12] C. Golgi, *Sulla fina anatomia degli organi centrali del sistema nervoso*. S. Calderini, 1885.

[13] S. Cajal, "Y, les nouvelles idées sur la fine anatomie des centres nerveux," 1894.

[14] A. Luchicchi, "Amsterdam vumc - lecturer luchicchi - drawing gross and microscopic areas and pathways in the human cortex - english labels." https://anatomytool.org/content/amsterdam-vumc-lecturer-luchicchi-drawing-gross-and-microscopic-areas-and-pathways-human-cortex-english-labels. Accessed: 2024-03-08.

[15] L. Pang, "Hallucinations experienced by visually impaired: Charles bonnet syndrome," *Optometry and Vision Science*, vol. 93, no. 12, pp. 1466–1478, 2016.

[16] R. Portillo-Lara, B. Tahirbegi, C. A. Chapman, J. A. Goding, and R. A. Green, "Mind the gap: State-of-the-art technologies and applications for eeg-based brain–computer interfaces," *APL bioengineering*, vol. 5, no. 3, 2021.

[17] J. Williams, G. Hek, A. Vardy, V. Rottschäfer, J. B. van den Berg, and J. Hulshof, "Mathematical techniques for neuromuscular analysis," *Proceedings of the 52nd European Study Group with Industry*, 2005.

[18] Bright Brain Centre, "Eeg and brainwaves." https://www.brightbraincentre.co.uk/electroencephalogram-eeg-brainwaves/, May 2021. Accessed: November 13, 2024.

[19] Y. Choe, "Processing of analogy in the thalamocortical circuit," in *Proceedings of the International Joint Conference on Neural Networks, 2003.*, vol. 2, pp. 1480–1485, IEEE, 2003.

[20] G. Ambrosi, D. Cantino, P. Castano, S. Correr, L. D'Este, R. Donato, G. Familiari, F. Fornai, M. Gulisano, A. Iannello, *et al.*, "Anatomia dell'uomo," *Seconda Edizione edi-ermes*, 2006.

[21] A. S. Asan, J. R. McIntosh, and J. B. Carmel, "Targeting sensory and motor integration for recovery of movement after cns injury," *Frontiers in Neuroscience*, vol. 15, p. 791824, 2022.

[22] "Homunculus | Description, History, Models, & Importance | Britannica — britannica.com." https://www.britannica.com/science/homunculus-biology. [Accessed 14-11-2024].

[23] G. Cheron, G. Petit, J. Cheron, A. Leroy, A. Cebolla, C. Cevallos, M. Petieau, T. Hoellinger, D. Zarka, A.-M. Clarinval, *et al.*, "Brain oscillations in sport: toward eeg biomarkers of performance," *Frontiers in psychology*, vol. 7, p. 246, 2016.

[24] R. Stickgold and M. P. Walker, *The neuroscience of sleep*. Academic Press, 2010.

[25] A. Rechtschaffen and A. Kales, "Proposed supplements and amendments to'a manual of standardized terminology, techniques and scoring system for sleep stages of human subjects', standard," *Psychiatry and Clinical Neurosciences*, vol. 55, pp. 305–310, 1968.

[26] H. L. Rusch, *Links between stress, sleep, and inflammation: a translational perspective of resilience*. Karolinska Institutet (Sweden), 2020.

[27] M. Troester, S. Quan, R. Berry, D. Plante, A. Abreu, and M. Alzoubaidi, "The aasm manual for the scoring of sleep and associated events: Rules, terminology and technical specifications," 2023.

[28] D. Moser, P. Anderer, G. Gruber, S. Parapatics, E. Loretz, M. Boeck, G. Kloesch, E. Heller, A. Schmidt, H. Danker-Hopfe, *et al.*, "Sleep classification according to aasm and rechtschaffen & kales: effects on sleep scoring parameters," *Sleep*, vol. 32, no. 2, pp. 139–149, 2009.

[29] J. J. Eggermont, "Chapter 10 - nonauditory effects of noise," in *Noise and the Brain* (J. J. Eggermont, ed.), pp. 266–300, San Diego: Academic Press, 2014.

[30] G. Curcio, M. Ferrara, and L. De Gennaro, "Sleep loss, learning capacity and academic performance," *Sleep medicine reviews*, vol. 10, no. 5, pp. 323–337, 2006.

[31] T. Åkerstedt, P. Philip, A. Capelli, and G. Kecklund, "Sleep loss and accidents—work hours, life style, and sleep pathology," *Progress in brain research*, vol. 190, pp. 169–188, 2011.

[32] M. S. Gazzaniga, R. B. Ivry, and G. Mangun, "Cognitive neuroscience. the biology of the mind,(2014)," 2006.

[33] D. Purves, R. Cabeza, S. A. Huettel, M. L. Platt, K. S. LaBar, and M. G. Woldorff, *Principles of cognitive neuroscience*. Sinauer Associates Sunderland, MA, 2013.

[34] P. B. Foley, *Encephalitis lethargica*. Springer, 2014.

[35] A. A. Borbély and P. Achermann, "Concepts and models of sleep regulation: an overview," *Journal of sleep research*, vol. 1, no. 2, pp. 63–79, 1992.

[36] T. Åkerstedt and S. Folkard, "The three-process model of alertness and its extension to performance, sleep latency, and sleep length," *Chronobiology international*, vol. 14, no. 2, pp. 115–123, 1997.

[37] C. Darwin, *The expression of the emotions in man and animals (1872)*. Penguin Books, 1993.

[38] P. Ekman, "An argument for basic emotions," *Cognition & emotion*, vol. 6, no. 3-4, pp. 169–200, 1992.

[39] R. Plutchik, "A general psychoevolutionary theory of emotion," *Emotion: Theory, research, and experience*, vol. 1, 1980.

[40] H. A. Hamzah and K. K. Abdalla, "Eeg-based emotion recognition systems; comprehensive study," *Heliyon*, vol. 10, no. 10, 2024.

[41] J. A. Russell, "A circumplex model of affect.," *Journal of Personality and Social Psychology*, vol. 39, no. 6, pp. 1161–1178, 1980.

[42] M. Tamietto and B. De Gelder, "Neural bases of the non-conscious perception of emotional signals," *Nature Reviews Neuroscience*, vol. 11, no. 10, pp. 697–709, 2010.

[43] "The experience of emotion." https://saylordotorg.github.io/text_introduction-to-psychology/s14-01-the-experience-of-emotion.html. [Accessed 14-11-2024].

[44] H. Barbas, S. Saha, N. Rempel-Clower, and T. Ghashghaei, "Serial pathways from primate prefrontal cortex to autonomic areas may influence emotional expression," *BMC neuroscience*, vol. 4, pp. 1–12, 2003.

[45] G. Pourtois, A. Schettino, and P. Vuilleumier, "Brain mechanisms for emotional influences on perception and attention: What is magic and what is not," *Biological psychology*, vol. 92, no. 3, pp. 492–512, 2013.

[46] A. Kujawa, D. N. Klein, and G. H. Proudfit, "Two-year stability of the late positive potential across middle childhood and adolescence," *Biological psychology*, vol. 94, no. 2, pp. 290–296, 2013.

[47] B.-W. Zhang, J. Xu, Y. Chang, H. Wang, H. Yao, and D. Tang, "Impaired cognitive reappraisal in panic disorder revealed by the late positive potential," *Neuroreport*, vol. 27, no. 2, pp. 99–103, 2016.

[48] J. S. Moser, G. Hajcak, E. Bukay, and R. F. Simons, "Intentional modulation of emotional responding to unpleasant pictures: An erp study," *Psychophysiology*, vol. 43, no. 3, pp. 292–296, 2006.

[49] H. T. Schupp, B. N. Cuthbert, M. M. Bradley, J. T. Cacioppo, T. Ito, and P. J. Lang, "Affective picture processing: the late positive potential is modulated by motivational relevance," *Psychophysiology*, vol. 37, no. 2, pp. 257–261, 2000.

[50] S. Horndasch, H. Heinrich, O. Kratz, and G. H. Moll, "The late positive potential as a marker of motivated attention to underweight bodies in girls with anorexia nervosa," *Journal of psychosomatic research*, vol. 73, no. 6, pp. 443–447, 2012.

[51] V. Mondini, *EEG-based Brain-Computer Interfaces for neurorehabilitation and control*. PhD thesis, University of Bologna, 2019.

[52] V. Mondini, "Sviluppo di un sistema bci eeg-based per il controllo del movimento di un modello di arto superiore," 2015.

[53] A. V. Oppenheim, *Discrete-time signal processing*. Pearson Education India, 1999.

[54] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[55] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006.

[56] K. Sacco *et al.*, *Le Neuroimmagini. Nuove frontiere per lo studio del cervello umano in vivo*. Idelson-Gnocchi, 2013.

[57] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[58] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[59] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[60] S. Weidman, *Deep learning from scratch: Building with python from first principles*. O'Reilly Media, 2019.

[61] R. Oostenveld, *Improving EEG source analysis using prior knowledge*. [Sl: sn], 2003.

[62] M. E. E. Alahi, Y. Liu, Z. Xu, H. Wang, T. Wu, and S. C. Mukhopadhyay, "Recent advancement of electrocorticography (ecog) electrodes for chronic neural recording/stimulation," *Materials Today Communications*, vol. 29, p. 102853, 2021.

[63] S. S. Balaji and K. K. Parhi, "Seizure onset zone identification from ieeg: A review," *IEEE Access*, vol. 10, pp. 62535–62547, 2022.

[64] J. A. Kim and K. D. Davis, "Magnetoencephalography: Physics, techniques, and applications in the basic and clinical neurosciences," *Journal of Neurophysiology*, vol. 125, no. 3, pp. 938–956, 2021.

[65] S. Baillet, "Magnetoencephalography for brain electrophysiology and imaging," *Nature neuroscience*, vol. 20, no. 3, pp. 327–339, 2017.

[66] B. D. Edmonds, W. Welch, Y. Sogawa, J. Mountz, A. Bagić, and C. Patterson, "The role of magnetoencephalography and single-photon emission computed tomography in evaluation of children with drug-resistant epilepsy," *Journal of Child Neurology*, vol. 36, no. 8, pp. 673–679, 2021.

[67] T. Kida, E. Tanaka, and R. Kakigi, "Multi-dimensional dynamics of human electromagnetic brain activity," *Frontiers in human neuroscience*, vol. 9, p. 713, 2016.

[68] A. Colon, H. Ronner, P. Boon, and P. Ossenblok, "Evaluation of meg vs eeg after sleep deprivation in epilepsy," *Acta Neurologica Scandinavica*, vol. 135, no. 2, pp. 247–251, 2017.

[69] "Structural vs Functional MRI Scan | BioRender Science Templates — biorender.com." https://www.biorender.com/template/structural-vs-functional-mri-scan. [Accessed 28-11-2024].

[70] R. K. Almajidy, K. Mankodiya, M. Abtahi, and U. G. Hofmann, "A newcomer's guide to functional near infrared spectroscopy experiments," *IEEE Reviews in Biomedical Engineering*, vol. 13, pp. 292–308, 2019.

[71] M. Ferrari and V. Quaresima, "A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application," *Neuroimage*, vol. 63, no. 2, pp. 921–935, 2012.

[72] S. Cerutti, A. M. Bianchi, and G. Rizzo, "The role of technology in the integration between biomedical signals and bioimaging," *Istituto Lombardo-Accademia di Scienze e Lettere• Incontri di Studio*, pp. 15–28, 2002.

[73] R. W. Homan, J. Herman, and P. Purdy, "Cerebral location of international 10–20 system electrode placement," *Electroencephalography and clinical neurophysiology*, vol. 66, no. 4, pp. 376–382, 1987.

[74] D. Lachner-Piza, *Distinguishing between epileptic and physiologic high-frequency brain activity using automatically detected EEG patterns*. PhD thesis, Dissertation, Universität Freiburg, 2022, 2022.

[75] L. De Gennaro and M. Ferrara, "Sleep spindles: an overview," *Sleep medicine reviews*, vol. 7, no. 5, pp. 423–440, 2003.

[76] F. Beverina, G. Palmas, S. Silvoni, F. Piccione, S. Giove, *et al.*, "User adaptive bcis: Ssvep and p300 based interfaces.," *PsychNology J.*, vol. 1, no. 4, pp. 331–354, 2003.

[77] "Event Related Potentials application for EEG Brain-Computer Interface — brainlatam.com." https://brainlatam.com/blog/event-related-potentials-application-for-eeg-brain-computer-interface-803. [Accessed 19-11-2024].

[78] G. L. Wallstrom, R. E. Kass, A. Miller, J. F. Cohn, and N. A. Fox, "Automatic correction of ocular artifacts in the eeg: a comparison of regression-based and component-based methods," *International journal of psychophysiology*, vol. 53, no. 2, pp. 105–119, 2004.

[79] C. N. Karson, R. S. Burns, P. A. LeWitt, N. L. Foster, and R. P. Newman, "Blink rates and disorders of movement," *Neurology*, vol. 34, no. 5, pp. 677–677, 1984.

[80] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from eeg signals: a review," *Sensors*, vol. 19, no. 5, p. 987, 2019.

[81] A. Q. Hamal and A. W. bin Abdul Rehman, "Artifact processing of epileptic eeg signals: an overview of different types of artifacts," in *2013 International Conference on Advanced Computer Science Applications and Technologies*, pp. 358–361, IEEE, 2013.

[82] K. J. Lee, C. Park, and B. Lee, "Elimination of ecg artifacts from a single-channel eeg using sparse derivative method," in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2384–2389, IEEE, 2015.

[83] G. Placidi, L. Cinque, and M. Polsinelli, "A fast and scalable framework for automated artifact recognition from eeg signals represented in scalp topographies of independent components," *Computers in Biology and Medicine*, vol. 132, p. 104347, 2021.

[84] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals—practices and possibilities," *IEEE transactions on information technology in biomedicine*, vol. 16, no. 3, pp. 488–500, 2012.

[85] L. Pion-Tonachini, K. Kreutz-Delgado, and S. Makeig, "Iclabel: An automated electroencephalographic independent component classifier, dataset, and website," *NeuroImage*, vol. 198, pp. 181–197, 2019.

[86] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "Bci2000: a general-purpose brain-computer interface (bci) system," *IEEE Transactions on biomedical engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

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

[88] A. Cicone, "Nonstationary signal decomposition for dummies," *Advances in Mathematical Methods and High Performance Computing*, pp. 69–82, 2019.

[89] D. Laszuk, "Python implementation of empirical mode decomposition algorithm." `https://github.com/laszukdawid/PyEMD`, 2017.

[90] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in adaptive data analysis*, vol. 1, no. 01, pp. 1–41, 2009.

[91] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 466, no. 2117, pp. 1291–1302, 2010.

[92] A. Cicone and H. Zhou, "Fast iterative filtering: a new, fast and robust decomposition method for nonlinear and nonstationary processes," in *EGU General Assembly Conference Abstracts*, pp. EGU21–3303, 2021.

[93] A. Cicone and E. Pellegrino, "Multivariate fast iterative filtering for the decomposition of nonstationary signals," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1521–1531, 2022.

[94] A. Cicone and H. Zhou, "Multidimensional iterative filtering method for the decomposition of high–dimensional non–stationary signals," *Numerical Mathematics: Theory, Methods and Applications*, vol. 10, no. 2, pp. 278–298, 2017.

[95] A. Cicone, G. D'Angelo, E. Papini, D. Recchiuti, and P. Ubertini, "The imfogram: a new time-frequency representation method for nonstationary signals analysis," in *AGU Fall Meeting Abstracts*, vol. 2022, pp. NG35B–0451, 2022.

[96] A. Cicone, J. Liu, and H. Zhou, "Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis," *Applied and Computational Harmonic Analysis*, vol. 41, no. 2, pp. 384–411, 2016.

[97] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis; using physiological signals," *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2011.

[98] Z. J. Koles, M. S. Lazar, and S. Z. Zhou, "Spatial patterns underlying population differences in the background eeg," *Brain topography*, vol. 2, no. 4, pp. 275–284, 1990.

[99] P. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[100] S. Makeig and R. Galambos, "The cerp: Event-related perturbations in steady-state responses," in *Brain dynamics: Progress and perspectives*, pp. 375–400, Springer, 1989.

[101] F. H. Duffy, J. L. Burchfiel, and C. T. Lombroso, "Brain electrical activity mapping (beam): A method for extending the clinical utility of eeg and evoked potential data," *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, vol. 5, no. 4, pp. 309–321, 1979.

[102] S.-E. Moon, C.-J. Chen, C.-J. Hsieh, J.-L. Wang, and J.-S. Lee, "Emotional eeg classification using connectivity features and convolutional neural networks," *Neural Networks*, vol. 132, pp. 96–107, 2020.

[103] M. X. Cohen, *Analyzing neural time series data: theory and practice*. MIT press, 2014.

[104] W. Liu, C. Zhang, X. Wang, J. Xu, Y. Chang, T. Ristaniemi, and F. Cong, "Functional connectivity of major depression disorder using ongoing eeg during music perception," *Clinical Neurophysiology*, vol. 131, no. 10, pp. 2413–2422, 2020.

[105] B. He, L. Astolfi, P. A. Valdés-Sosa, D. Marinazzo, S. O. Palva, C.-G. Bénar, C. M. Michel, and T. Koenig, "Electrophysiological brain connectivity: theory and implementation," *IEEE transactions on biomedical engineering*, vol. 66, no. 7, pp. 2115–2137, 2019.

[106] C. J. Stam, G. Nolte, and A. Daffertshofer, "Phase lag index: assessment of functional connectivity from multi channel eeg and meg with diminished bias from common sources," *Human brain mapping*, vol. 28, no. 11, pp. 1178–1193, 2007.

[107] G. C. O'Neill, P. Tewarie, D. Vidaurre, L. Liuzzi, M. W. Woolrich, and M. J. Brookes, "Dynamics of large-scale electrophysiological networks: A technical review," *Neuroimage*, vol. 180, pp. 559–576, 2018.

[108] B. Zhang, G. Yan, Z. Yang, Y. Su, J. Wang, and T. Lei, "Brain functional networks based on resting-state eeg data for major depressive disorder analysis and classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 215–229, 2021.

[109] C. Brunner, M. Billinger, M. Seeber, T. R. Mullen, and S. Makeig, "Volume conduction influences scalp-based connectivity estimates," *Frontiers in computational neuroscience*, vol. 10, p. 121, 2016.

[110] J. Wu, J. Zhang, C. Liu, D. Liu, X. Ding, and C. Zhou, "Graph theoretical analysis of eeg functional connectivity during music perception," *Brain research*, vol. 1483, pp. 71–81, 2012.

[111] S. J. Ruiz-Gómez, R. Hornero, J. Poza, A. Maturana-Candelas, N. Pinto, and C. Gómez, "Computational modeling of the effects of eeg volume conduction on functional connectivity metrics. application to alzheimer's disease continuum," *Journal of neural engineering*, vol. 16, no. 6, p. 066019, 2019.

[112] A. M. Maitin, A. Nogales, P. Chazarra, and Á. J. García-Tejedor, "Eegraph: An open-source python library for modeling electroencephalograms using graphs," *Neurocomputing*, vol. 519, pp. 127–134, 2023.

[113] A. Bradley, J. Yao, J. Dewald, and C.-P. Richter, "Evaluation of electroencephalography source localization algorithms with multiple cortical sources," *PloS one*, vol. 11, no. 1, p. e0147266, 2016.

[114] H. Sheikh, C. Prins, and E. Schrijvers, "Artificial intelligence: definition and background," in *Mission AI: The new system technology*, pp. 15–41, Springer, 2023.

[115] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, no. 4, pp. 303–314, 1989.

[116] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural networks*, vol. 4, no. 2, pp. 251–257, 1991.

[117] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.

[118] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain.," *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[119] S. Lahiri and K. Ghanta, "Artificial neural network model with the parameter tuning assisted by a differential evolution technique: The study of the hold up of the slurry flow in a pipeline," *Chemical Industry and Chemical Engineering Quarterly*, vol. 15, no. 2, pp. 103–117, 2009.

[120] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proceedings of 2010 IEEE international symposium on circuits and systems*, pp. 253–256, IEEE, 2010.

[121] D. H. Hubel, T. N. Wiesel, *et al.*, "Receptive fields of single neurones in the cat's striate cortex," *J physiol*, vol. 148, no. 3, pp. 574–591, 1959.

[122] N. I. Alonso *et al.*, "The mathematics of recurrent neural networks," *The Mathematics of Recurrent Neural Networks (October 27, 2024)*, 2024.

[123] C. Shi, Z. Zhang, W. Zhang, C. Zhang, and Q. Xu, "Learning multiscale temporal–spatial–spectral features via a multipath convolutional lstm neural network for change detection with hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2022.

[124] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 15, no. 5, p. 056013, 2018.

[125] H. Gou, Y. Piao, J. Ren, Q. Zhao, Y. Chen, C. Liu, W. Hong, and X. Zhang, "A solution to supervised motor imagery task in the bci controlled robot contest in world robot contest," *Brain Science Advances*, vol. 8, no. 2, pp. 153–161, 2022.

[126] H. Raza, A. Chowdhury, and S. Bhattacharyya, "Deep learning based prediction of eeg motor imagery of stroke patients' for neuro-rehabilitation application," in *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2020.

[127] A. Salami, J. Andreu-Perez, and H. Gillmeister, "Eeg-itnet: An explainable inception temporal convolutional network for motor imagery classification," *IEEE Access*, vol. 10, pp. 36672–36685, 2022.

[128] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[129] L. D. Fiederer, M. Völker, R. T. Schirrmeister, W. Burgard, J. Boedecker, and T. Ball, "Hybrid brain-computer-interfacing for human-compliant robots: inferring continuous subjective ratings with deep regression," *Frontiers in Neurorobotics*, vol. 13, p. 76, 2019.

[130] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An eeg-based brain-computer interface for cursor control," *Electroencephalography and clinical neurophysiology*, vol. 78, no. 3, pp. 252–259, 1991.

[131] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[132] S. Saha, K. A. Mamun, K. Ahmed, R. Mostafa, G. R. Naik, S. Darvishi, A. H. Khandoker, and M. Baumert, "Progress in brain computer interface: Challenges and opportunities," *Frontiers in systems neuroscience*, vol. 15, p. 578875, 2021.

[133] J. N. Mak and J. R. Wolpaw, "Clinical applications of brain-computer interfaces: current state and future prospects," *IEEE reviews in biomedical engineering*, vol. 2, pp. 187–199, 2009.

[134] L. Angrisani, P. Arpaia, A. Esposito, L. Gargiulo, A. Natalizio, G. Mastrati, N. Moccaldi, and M. Parvis, "Passive and active brain-computer interfaces for rehabilitation in health 4.0," *Measurement: Sensors*, vol. 18, p. 100246, 2021.

[135] I. Käthner, S. C. Wriessnegger, G. R. Müller-Putz, A. Kübler, and S. Halder, "Effects of mental workload and fatigue on the p300, alpha and theta band power during operation of an erp (p300) brain–computer interface," *Biological psychology*, vol. 102, pp. 118–129, 2014.

[136] G. R. Müller-Putz, A. Schwarz, J. Pereira, and P. Ofner, "From classic motor imagery to complex movement intention decoding: the noninvasive graz-bci approach," *Progress in brain research*, vol. 228, pp. 39–70, 2016.

[137] G. Placidi, M. Polsinelli, M. Spezialetti, and L. Cinque, "Bci driven by self-induced emotions: a multi-class study," in *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 1–6, IEEE, 2018.

[138] Y. Qin, Y. Zhang, Y. Zhang, S. Liu, and X. Guo, "Application and development of eeg acquisition and feedback technology: A review," *Biosensors*, vol. 13, no. 10, p. 930, 2023.

[139] "a. Filtering — eeglab.org." https://eeglab.org/tutorials/05_Preprocess/Filtering.html. [Accessed 26-11-2024].

[140] A. Delorme, "Eeg is better left alone," *Scientific reports*, vol. 13, no. 1, p. 2372, 2023.

[141] M. Schmoigl-Tonis, C. Schranz, and G. R. Müller-Putz, "Methods for motion artifact reduction in online brain-computer interface experiments: a systematic review," *Frontiers in Human Neuroscience*, vol. 17, p. 1251690, 2023.

[142] N. Bigdely-Shamlo, T. Mullen, C. Kothe, K.-M. Su, and K. A. Robbins, "The prep pipeline: standardized preprocessing for large-scale eeg analysis," *Frontiers in neuroinformatics*, vol. 9, p. 16, 2015.

[143] S. S. Al-Fraiji and D. Al-Shammary, "Survey for electroencephalography eeg signal classification approaches," in *Mobile Computing and Sustainable Informatics: Proceedings of ICMCSI 2021*, pp. 199–214, Springer, 2022.

[144] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for eeg-based brain–computer interfaces: a 10 year update," *Journal of neural engineering*, vol. 15, no. 3, p. 031005, 2018.

[145] G. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller, "Better than random: a closer look on bci results," *International journal of bioelectromagnetism*, vol. 10, no. 1, pp. 52–55, 2008.

[146] E. Theodoridou, L. Cinque, F. Mignosi, G. Placidi, M. Polsinelli, J. M. R. Tavares, and M. Spezialetti, "Hand tracking and gesture recognition by multiple contactless sensors: a survey," *IEEE Transactions on Human-Machine Systems*, 2022.

[147] G. Placidi, D. Avola, D. Iacoviello, and L. Cinque, "Overall design and implementation of the virtual glove," *Computers in biology and medicine*, vol. 43, no. 11, pp. 1927–1940, 2013.

[148] M. Polsinelli, "Progettazione, implementazione e caratterizzazione di un guanto virtuale basato su due leap motion controllers ortogonali," tesi di laurea, Università degli studi dell'Aquila, Dipartimento di ingegneria e scienze dell'informazione e matematica, L'Aquila, 2018. In testa al front.: Universita' degli studi dell'Aquila, Dipartimento di ingegneria e scienze dell'informazione e matematica, Corso di laurea in ingegneria informatica e automatica. - A.a. 2017-2018.

[149] A. DI MATTEO, "Sviluppo di un'interfaccia per l'integrazione dei dati da virtual glove in ambiente unity," *Università degli Studi del'Aquila*, 2019.

[150] G. Placidi, L. Cinque, M. Polsinelli, and M. Spezialetti, "Measurements by a leap-based virtual glove for the hand rehabilitation," *Sensors*, vol. 18, no. 3, p. 834, 2018.

[151] G. Placidi, D. Avola, L. Cinque, M. Polsinelli, E. Theodoridou, and J. M. R. Tavares, "Data integration by two-sensors in a leap-based virtual glove for human-system interaction," *Multimedia Tools and Applications*, vol. 80, no. 12, pp. 18263–18277, 2021.

[152] D. Gwon, K. Won, M. Song, C. S. Nam, S. C. Jun, and M. Ahn, "Review of public motor imagery and execution datasets in brain-computer interfaces," *Frontiers in Human Neuroscience*, vol. 17, p. 1134869, 2023.

[153] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "Bci2000: a general-purpose brain-computer interface (bci) system," *IEEE Transactions on biomedical engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[154] M. Wairagkar, "Eeg data for voluntary finger tapping movement," *University of Reading*, 2017.

[155] O. Ali, M. Saif-ur Rehman, T. Glasmachers, I. Iossifidis, and C. Klaes, "Contranet: A hybrid network for improving the classification of eeg and emg signals with limited training data," *Computers in Biology and Medicine*, vol. 168, p. 107649, 2024.

[156] P. Wang, M. Wang, Y. Zhou, Z. Xu, and D. Zhang, "Multiband decomposition and spectral discriminative analysis for motor imagery bci via deep neural network," *Frontiers of Computer Science*, vol. 16, pp. 1–13, 2022.

[157] C. Fan, B. Yang, X. Li, and P. Zan, "Temporal-frequency-phase feature classification using 3d-convolutional neural networks for motor imagery and movement," *Frontiers in Neuroscience*, vol. 17, p. 1250991, 2023.

[158] B. Taşar and O. Yaman, "Eeg signals based motor imagery and movement classification for bci applications," in *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, pp. 1425–1429, 2022.

[159] X. Wang, M. Hersche, B. Tömekce, B. Kaya, M. Magno, and L. Benini, "An accurate eegnet-based motor-imagery brain–computer interface for low-power edge computing," in *2020 IEEE international symposium on medical measurements and applications (MeMeA)*, pp. 1–6, IEEE, 2020.

[160] W. Fadel, C. Kollod, M. Wahdow, Y. Ibrahim, and I. Ulbert, "Multi-class classification of motor imagery eeg signals using image-based deep recurrent convolutional neural network," in *2020 8th International Winter Conference on Brain-Computer Interface (BCI)*, pp. 1–4, IEEE, 2020.

[161] F. Bencivenga, V. Sulpizio, M. G. Tullo, and G. Galati, "Assessing the effective connectivity of premotor areas during real vs imagined grasping: a dcm-peb approach," *Neuroimage*, vol. 230, p. 117806, 2021.

[162] A. M. Batula, J. A. Mark, Y. E. Kim, and H. Ayaz, "Comparison of brain activation during motor imagery and motor movement using fnirs," *Computational intelligence and neuroscience*, vol. 2017, 2017.

[163] P. Ofner, A. Schwarz, J. Pereira, and G. R. Müller-Putz, "Upper limb movements can be decoded from the time-domain of low-frequency eeg," *PloS one*, vol. 12, no. 8, p. e0182578, 2017.

[164] J. Mwata-Velu, T.; Ruiz-Pinales, H. Rostro-Gonzalez, M. Ibarra-Manzano, J. Cruz-Duarte, and J. Avina-Cervantes, "Motor imagery classification based on a recurrent-convolutional architecture to control a hexapod robot.," *Mathematics*, vol. 9, 2023.

[165] B. Blankertz, K.-R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. Millán, M. Schröder, and N. Birbaumer, "The bci competition iii: Validating alternative approaches to actual bci problems," *IEEE Transaction on Neural System and Rehabilitation Engineering*, vol. 14, no. 2, 2006.

[166] G. Dornhege, B. Blankertz, G. Curio, and K.-R. Müller, "Boosting bit rates in noninvasive eeg single-trial classifications by feature combination and multiclass paradigms," *IEEE Transactions on biomedical engineering*, vol. 51, no. 6, 2004.

[167] C. Brunner, R. Leeb, and G. Müller-Putz, "Bci competition 2008–graz data set a," *IEEE Dataport*, 2024.

[168] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, "Brain–computer communication: Motivation, aim, and impact of exploring a virtual apartment," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 6, 2007.

[169] H. Alwasiti, M. Z. Yusoff, and K. Raza, "Motor imagery classification for brain computer interface using deep metric learning," *IEEE Access*, vol. 8, 2020.

[170] M. Xu, J. Yao, Z. Zhang, R. Li, B. Yang, C. Li, J. Li, and J. Zhang, "Learning eeg topographical representation for classification via convolutional neural network," *Pattern Recognition*, vol. 105, 2020.

[171] D. Zhang, K. Chen, D. Jian, and L. Yao, "Motor imagery classification via temporal attention cues of graph embedded eeg signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, 2020.

[172] D. Lozzi, F. Mignosi, M. Spezialetti, G. Placidi, and M. Polsinelli, "A 4d lstm network for emotion recognition from the cross-correlation of the power spectral density of eeg signals," in *2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pp. 652–657, IEEE, 2022.

[173] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, "Emotion recognition based on eeg using lstm recurrent neural network," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 10, 2017.

[174] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Hum. Brain Mapp*, vol. 38, pp. 5391–5420, 2017.

[175] G. Placidi, G. De Gasperis, F. Mignosi, M. Polsinelli, and M. Spezialetti, "Integration of a bci with a hand tracking system and a motorized robotic arm to improve decoding of brain signals related to hand and finger movements," in *International Symposium on Visual Computing*, pp. 305–315, Springer, 2021.

[176] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 15, 2018.

[177] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. S. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013.

[178] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," in *NIPS-W*, 2017.

[179] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human Brain Mapping*, aug 2017.

[180] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-r. Muller, "Optimizing spatial filters for robust eeg single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41–56, 2008.

[181] S. L. Smith, P.-J. Kindermans, C. Ying, and Q. V. Le, "Don't decay the learning rate, increase the batch size," *arXiv preprint arXiv:1711.00489*, 2017.

[182] P. Di Giamberardino, D. Iacoviello, G. Placidi, M. Polsinelli, and M. Spezialetti, "A brain computer interface by eeg signals from self-induced emotions," in *VipIMAGE 2017: Proceedings of the VI ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing Porto, Portugal, October 18-20, 2017*, pp. 713–721, Springer, 2018.

[183] D. Lozzi, F. Mignosi, G. Placidi, and M. Polsinelli, "Graph model of phase lag index for connectivity analysis in eeg of emotions," in *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 348–353, IEEE, 2023.

[184] A. Erfanian and A. Erfani, "Ica-based classification scheme for eeg-based brain-computer interface: the role of mental practice and concentration skills," in *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 1, pp. 235–238, IEEE, 2004.

[185] D. Jaipriya and K. Sriharipriya, "Brain computer interface-based signal processing techniques for feature extraction and classification of motor imagery using eeg: A literature review," *Biomedical Materials & Devices*, pp. 1–13, 2023.

[186] A. Al-Saegh, S. A. Dawwd, and J. M. Abdul-Jabbar, "Deep learning for motor imagery eeg-based classification: A review," *Biomedical Signal Processing and Control*, vol. 63, p. 102172, 2021.

[187] N. Tibrewal, N. Leeuwis, and M. Alimardani, "Classification of motor imagery eeg using deep learning increases performance in inefficient bci users," *Plos one*, vol. 17, no. 7, p. e0268880, 2022.

[188] G. R. Müller-Putz, R. J. Kobler, J. Pereira, C. Lopes-Dias, L. Hehenberger, V. Mondini, V. Martínez-Cagigal, N. Srisrisawang, H. Pulferer, L. Batistić, *et al.*, "Feel your reach: An eeg-based framework to continuously detect goal-directed movements and error processing to gate kinesthetic feedback informed artificial arm control," *Frontiers in Human Neuroscience*, vol. 16, p. 841312, 2022.

[189] D. Borra, S. Fantozzi, and E. Magosso, "Eeg motor execution decoding via interpretable sinc-convolutional neural networks," in *Mediterranean Conference on Medical and Biological Engineering and Computing*, pp. 1113–1122, Springer, 2019.

[190] A. Schwarz, M. K. Höller, J. Pereira, P. Ofner, and G. R. Müller-Putz, "Decoding hand movements from human eeg to control a robotic arm in a simulation environment," *Journal of neural engineering*, vol. 17, no. 3, p. 036010, 2020.

[191] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human Brain Mapping*, aug 2017.

[192] S. Makeig, A. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Advances in neural information processing systems*, vol. 8, 1995.

[193] A. Delorme and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.

[194] V. Jurcak, D. Tsuzuki, and I. Dan, "10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems," *Neuroimage*, vol. 34, no. 4, pp. 1600–1611, 2007.

[195] A. Jiang, J. Shang, X. Liu, Y. Tang, H. K. Kwan, and Y. Zhu, "Efficient csp algorithm with spatio-temporal filtering for motor imagery classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 4, pp. 1006–1016, 2020.

[196] T. Viering and M. Loog, "The shape of learning curves: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 6, pp. 7799–7819, 2023.

[197] R. E. Elmahdi, N. Y. Ahmed, M. B. M. Amin, and A. O. Hamza, "A review of using common spatial pattern in discrimination of multiclass motor imagery–based brain-computer interface," *Journal of Clinical Engineering*, vol. 45, no. 1, pp. 45–50, 2020.

[198] A. Pinegger, J. Faller, S. Halder, S. C. Wriessnegger, and G. R. Müller-Putz, "Control or non-control state: that is the question! an asynchronous visual p300-based bci approach," *Journal of neural engineering*, vol. 12, no. 1, p. 014001, 2015.

[199] C.-I. Hung, P.-L. Lee, Y.-T. Wu, L.-F. Chen, T.-C. Yeh, and J.-C. Hsieh, "Recognition of motor imagery electroencephalography using independent component analysis and machine classifiers," *Annals of biomedical engineering*, vol. 33, pp. 1053–1070, 2005.

[200] M. G. Antonelli, P. Beomonte Zobel, C. Manes, E. Mattei, and N. Stampone, "Emotional intelligence for the decision-making process of trajectories in collaborative robotics," *Machines*, vol. 12, no. 2, p. 113, 2024.

[201] G. Placidi, "A smart virtual glove for the hand telerehabilitation," *Computers in Biology and Medicine*, vol. 37, no. 8, pp. 1100–1107, 2007.

[202] G. Placidi, A. Di Matteo, D. Lozzi, M. Polsinelli, and E. Theodoridou, "Patient–therapist cooperative hand telerehabilitation through a novel framework involving the virtual glove system," *Sensors*, vol. 23, no. 7, p. 3463, 2023.

[203] A. I. Sburlea and G. R. Müller-Putz, "Exploring representations of human grasping in neural, muscle and kinematic signals," *Scientific reports*, vol. 8, no. 1, p. 16669, 2018.

[204] E. Mattei, D. Lozzi, A. Di Matteo, M. Polsinelli, C. Manes, F. Mignosi, and G. Placidi, "Deep learning architecture analysis for eeg-based bci classification under motor execution," in *2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 549, 555, IEEE, 2024.

[205] K. E. Mathewson, T. J. Harrison, and S. A. Kizuk, "High and dry? comparing active dry eeg electrodes to active and passive wet electrodes," *Psychophysiology*, vol. 54, no. 1, pp. 74–82, 2017.

[206] M. A. Lopez-Gordo, D. Sanchez-Morillo, and F. P. Valle, "Dry eeg electrodes," *Sensors*, vol. 14, no. 7, pp. 12847–12870, 2014.

[207] T. Schneider, X. Wang, M. Hersche, L. Cavigelli, and L. Benini, "Q-eegnet: An energy-efficient 8-bit quantized parallel eegnet implementation for edge motor-imagery brain-machine interfaces," *Proceedings of the IEEE International Conference on Smart Computing*, 2020.

[208] P. Aricò, G. Borghini, G. Di Flumeri, N. Sciaraffa, and F. Babiloni, "Passive bci beyond the lab: current trends and future directions," *Physiological measurement*, vol. 39, no. 8, p. 08TR02, 2018.

[209] D. Leong, "Ventral and dorsal stream eeg channels: Key features for eeg-based object recognition and identification," *Ieee Transactions on Neural Systems and Rehabilitation Engineering*, 2023.

[210] A. Di Matteo, D. Lozzi, F. Mignosi, M. Polsinelli, and G. Placidi, "A dicom-based standard for quantitative physical rehabilitation," *Computational and Structural Biotechnology Journal*, 2025.

[211] T.-W. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources," *Neural computation*, vol. 11, no. 2, pp. 417–441, 1999.

[212] D. Dharmaprani, H. K. Nguyen, T. W. Lewis, D. DeLosAngeles, J. O. Willoughby, and K. J. Pope, "A comparison of independent component analysis algorithms and measures to discriminate between eeg and artifact components," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 825–828, IEEE, 2016.

[213] M. Polsinelli, A. Di Matteo, D. Lozzi, E. Mattei, F. Mignosi, L. Nazzicone, V. Stornelli, and G. Placidi, "Portable head-mounted system for mobile forearm tracking," *Sensors*, vol. 24, no. 7, p. 2227, 2024.

[214] L. Pion-Tonachini, S.-H. Hsu, S. Makeig, T.-P. Jung, and G. Cauwenberghs, "Real-time eeg source-mapping toolbox (rest): Online ica and source localization," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 4114–4117, IEEE, 2015.

[215] H. Akbari, M. T. Sadiq, N. Jafari, J. Too, N. Mikaeilvand, A. Cicone, and S. Serra Capizzano, "Recognizing seizure using poincaré plot of eeg signals and graphical features in dwt domain," *Bratislava Medical Journal*, 2023.

[216] M. Aljalal, S. Ibrahim, R. Djemal, and W. Ko, "Comprehensive review on brain-controlled mobile robots and robotic arms based on electroencephalography signals," *Intelligent service robotics*, vol. 13, no. 4, pp. 539–563, 2020.

[217] A. Toichoa Eyam, W. M. Mohammed, and J. L. Martinez Lastra, "Emotion-driven analysis and control of human-robot interactions in collaborative applications," *Sensors*, vol. 21, no. 14, p. 4626, 2021.

[218] C. J. Hilditch and A. W. McHill, "Sleep inertia: current insights," *Nature and science of sleep*, pp. 155–165, 2019.

[219] P. Tassi and A. Muzet, "Sleep inertia," *Sleep medicine reviews*, vol. 4, no. 4, pp. 341–353, 2000.

[220] M. H. Kryger, *Principles and Practice of Sleep Medicine*. Philadelphia, PA: Elsevier, sixth ed., 2017.

[221] H. K. Ritchie, T. M. Burke, T. B. Dear, A. W. Mchill, J. Axelsson, and K. P. Wright Jr, "Impact of sleep inertia on visual selective attention for rare targets and the influence of chronotype," *Journal of sleep research*, vol. 26, no. 5, pp. 551–558, 2017.

[222] N. R. Carlson and N. R. Carlson, *Physiology of behavior*. Pearson Boston, 2007.

[223] T. J. Balkin, A. R. Braun, N. J. Wesensten, K. Jeffries, M. Varga, P. Baldwin, G. Belenky, and P. Herscovitch, "The process of awakening: a pet study of regional brain activity patterns mediating the re-establishment of alertness and consciousness," *Brain*, vol. 125, no. 10, pp. 2308–2319, 2002.

[224] E. A. Nofzinger, "What can neuroimaging findings tell us about sleep disorders?," *Sleep Medicine*, vol. 5, pp. S16–S22, 2004.

[225] K. H. Jawabri and A. Raja, "Physiology, sleep patterns," in *StatPearls [Internet]*, StatPearls Publishing, 2022.

[226] J. Wörle, B. Metz, M. B. Steinborn, L. Huestegge, and M. Baumann, "Differential effects of driver sleepiness and sleep inertia on driving behavior," *Transportation research part F: traffic psychology and behaviour*, vol. 82, pp. 111–120, 2021.

[227] B. C. Tefft *et al.*, *Prevalence of motor vehicle crashes involving drowsy drivers, United States, 2009-2013*. AAA Foundation for Traffic Safety Washington, DC, 2014.

[228] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, C. Marcus, B. V. Vaughn, *et al.*, "The aasm manual for the scoring of sleep and associated events," *Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine*, vol. 176, p. 2012, 2012.

[229] H. W. Loh, C. P. Ooi, J. Vicnesh, S. L. Oh, O. Faust, A. Gertych, and U. R. Acharya, "Automated detection of sleep stages using deep learning techniques: A systematic review of the last decade (2010–2020)," *Applied Sciences*, vol. 10, no. 24, p. 8963, 2020.

[230] K. A. I. Aboalayon, M. Faezipour, W. S. Almuhammadi, and S. Moslehpour, "Sleep stage classification using eeg signal analysis: a comprehensive survey and new investigation," *Entropy*, vol. 18, no. 9, p. 272, 2016.

[231] D. Sarkar, D. Guha, P. Tarafdar, S. Sarkar, A. Ghosh, and D. Dey, "A comprehensive evaluation of contemporary methods used for automatic sleep staging," *Biomedical Signal Processing and Control*, vol. 77, p. 103819, 2022.

[232] R. Boostani, F. Karimzadeh, and M. Nami, "A comparative review on sleep stage classification methods in patients and healthy individuals," *Computer methods and programs in biomedicine*, vol. 140, pp. 77–91, 2017.

[233] R. N. Sekkal, F. Bereksi-Reguig, D. Ruiz-Fernandez, N. Dib, and S. Sekkal, "Automatic sleep stage classification: From classical machine learning methods to deep learning," *Biomedical Signal Processing and Control*, vol. 77, p. 103751, 2022.

[234] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[235] B. Kemp, A. H. Zwinderman, B. Tuk, H. A. Kamphuisen, and J. J. Oberye, "Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the eeg," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 9, pp. 1185–1194, 2000.

[236] A. Supratak, H. Dong, C. Wu, and Y. Guo, "Deepsleepnet: A model for automatic sleep stage scoring based on raw single-channel eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 1998–2008, 2017.

[237] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (eeg) classification tasks: a review," *Journal of neural engineering*, vol. 16, no. 3, p. 031001, 2019.

[238] J. Zhang and Y. Wu, "Complex-valued unsupervised convolutional neural networks for sleep stage classification," *Computer methods and programs in biomedicine*, vol. 164, pp. 181–191, 2018.

[239] A. Vilamala, K. H. Madsen, and L. K. Hansen, "Deep convolutional neural networks for interpretable analysis of eeg sleep stage scoring," in *2017 IEEE 27th international workshop on machine learning for signal processing (MLSP)*, pp. 1–6, IEEE, 2017.

[240] A. H. Ansari, O. De Wel, M. Lavanga, A. Caicedo, A. Dereymaeker, K. Jansen, J. Vervisch, M. De Vos, G. Naulaers, and S. Van Huffel, "Quiet sleep detection in preterm infants using deep convolutional neural networks," *Journal of neural engineering*, vol. 15, no. 6, p. 066006, 2018.

[241] E. Bresch, U. Großekathöfer, and G. Garcia-Molina, "Recurrent deep neural networks for real-time sleep stage classification from single channel eeg," *Frontiers in computational neuroscience*, vol. 12, p. 85, 2018.

[242] A. Supratak, H. Dong, C. Wu, and Y. Guo, "Deepsleepnet: A model for automatic sleep stage scoring based on raw single-channel eeg," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 1998–2008, 2017.

[243] H. Dong, A. Supratak, W. Pan, C. Wu, P. M. Matthews, and Y. Guo, "Mixed neural network approach for temporal sleep stage classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 2, pp. 324–333, 2017.

[244] M. Shahin, B. Ahmed, S. T.-B. Hamida, F. L. Mulaffer, M. Glos, and T. Penzel, "Deep learning and insomnia: assisting clinicians with their diagnosis," *IEEE journal of biomedical and health informatics*, vol. 21, no. 6, pp. 1546–1553, 2017.

[245] O. Tsinalis, P. M. Matthews, and Y. Guo, "Automatic sleep stage scoring using time-frequency analysis and stacked sparse autoencoders," *Annals of biomedical engineering*, vol. 44, pp. 1587–1597, 2016.

[246] C.-E. Kuo and G.-T. Chen, "Automatic sleep staging based on a hybrid stacked lstm neural network: verification using large-scale dataset," *IEEE access*, vol. 8, pp. 111837–111849, 2020.

[247] E. Alickovic and A. Subasi, "Ensemble svm method for automatic sleep stage classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 6, pp. 1258–1265, 2018.

[248] S. A. Imtiaz and E. Rodriguez-Villegas, "Automatic sleep staging using state machine-controlled decision trees," in *2015 37th Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pp. 378–381, IEEE, 2015.

[249] I. Choi and W. Sung, "Sleep model–a sequence model for predicting the next sleep stage," *arXiv preprint arXiv:2302.12709*, 2023.

[250] K. Slyusarenko and I. Fedorin, "Smart alarm based on sleep stages prediction," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 4286–4289, IEEE, 2020.

[251] D. Schwarze, F. Diederichs, L. Weiser, H. Widlroither, R. Verhoeven, and M. Rötting, "Are drivers allowed to sleep? sleep inertia effects drivers' performance after different sleep durations in automated driving," *Multimodal Technologies and Interaction*, vol. 7, no. 6, p. 62, 2023.

[252] G. Romyn, G. D. Roach, M. Lastella, D. J. Miller, N. G. Versey, and C. Sargent, "The impact of sleep inertia on physical, cognitive, and subjective performance following a 1-or 2-hour afternoon nap in semiprofessional athletes," *International Journal of Sports Physiology and Performance*, vol. 17, no. 7, pp. 1140–1150, 2022.

[253] D. Bruck and D. L. Pisani, "The effects of sleep inertia on decision-making performance," *Journal of sleep research*, vol. 8, no. 2, pp. 95–103, 1999.

[254] M. E. Jewett, J. K. Wyatt, A. RITZ-DE CECCO, S. B. Khalsa, D.-J. DIJK, and C. A. Czeisler, "Time course of sleep inertia dissipation in human performance and alertness," *Journal of sleep research*, vol. 8, no. 1, pp. 1–8, 1999.

[255] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, *et al.*, "Meg and eeg data analysis with mne-python," *Frontiers in neuroscience*, p. 267, 2013.

[256] D. J. Dijk, D. G. Beersma, and G. M. Bloem, "Sex differences in the sleep eeg of young adults: visual scoring and spectral analysis," *Sleep*, vol. 12, no. 6, pp. 500–507, 1989.

[257] J. Favaro, M. A. Colombo, E. Mikulan, S. Sartori, M. Nosadini, M. F. Pelizza, M. Rosanova, S. Sarasso, M. Massimini, and I. Toldo, "The maturation of aperiodic eeg activity across development reveals a progressive differentiation of wakefulness from sleep," *NeuroImage*, p. 120264, 2023.

[258] M. Mourtazaev, B. Kemp, A. Zwinderman, and H. Kamphuisen, "Age and gender affect different characteristics of slow waves in the sleep eeg," *Sleep*, vol. 18, no. 7, pp. 557–564, 1995.

[259] L. R. Medsker and L. Jain, "Recurrent neural networks," *Design and Applications*, vol. 5, pp. 64–67, 2001.

[260] C. Centofanti, W. Tiberti, A. Marotta, F. Graziosi, and D. Cassioli, "Taming latency at the edge: A user-aware service placement approach," *Computer Networks*, vol. 247, p. 110444, 2024.

[261] G. Curcio, M. Casagrande, and M. Bertini, "Sleepiness: evaluating and quantifying methods," *International Journal of Psychophysiology*, vol. 41, no. 3, pp. 251–263, 2001.

[262] J. Su, J. Zhu, T. Song, and H. Chang, "Subject-independent eeg emotion recognition based on genetically optimized projection dictionary pair learning," *Brain Sciences*, vol. 13, no. 7, p. 977, 2023.

[263] R. Liu, Y. Chao, X. Ma, X. Sha, L. Sun, S. Li, and S. Chang, "Ertnet: an interpretable transformer-based framework for eeg emotion recognition," *Frontiers in Neuroscience*, vol. 18, p. 1320645, 2024.

[264] M. Spezialetti, G. Placidi, and S. Rossi, "Emotion recognition for human-robot interaction: Recent advances and future perspectives," *Frontiers in Robotics and AI*, vol. 7, p. 532279, 2020.

[265] D. Lozzi, E. Mattei, R. Ciuffini, A. Di Matteo, A. Marrelli, R. Ornello, M. Polsinelli, C. Rosignoli, S. Sacco, and G. Placidi, "The challenge of driving bci with emotional signals collected by eeg," in *9th Graz Brain-Computer Interface Conference 2024*, pp. 366–371, 2024.

[266] E. P. Torres, E. A. Torres, M. Hernández-Álvarez, and S. G. Yoo, "Eeg-based bci emotion recognition: A survey," *Sensors*, vol. 20, no. 18, p. 5083, 2020.

[267] P. R. Yeganeh, A. Nouri, T. Khatibi, and A. Khanteymoori, "Emotion recognition from electroencephalogram signals based on deep neural networks," *International Journal of Bioinformatics and Intelligent Computing*, vol. 2, no. 2, pp. 212–230, 2023.

[268] O. Bazgir, Z. Mohammadi, and S. A. H. Habibi, "Emotion recognition with machine learning using eeg signals," in *2018 25th national and 3rd international iranian conference on biomedical engineering (ICBME)*, pp. 1–5, IEEE, 2018.

[269] M. A. Asghar, M. J. Khan, x. Fawad, Y. Amin, M. Rizwan, M. Rahman, S. Badnava, and S. S. Mirjavadi, "Eeg-based multi-modal emotion recognition using bag of deep features: An optimal feature selection approach," *Sensors*, vol. 19, no. 23, p. 5218, 2019.

[270] G. Placidi, M. Polsinelli, M. Spezialetti, L. Cinque, P. Di Giamberardino, and D. Iacoviello, "Self-induced emotions as alternative paradigm for driving brain–computer interfaces," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 2018.

[271] G. Placidi, A. Petracca, M. Spezialetti, and D. Iacoviello, "Classification strategies for a single-trial binary brain computer interface based on remembering unpleasant odors," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 7019–7022, IEEE, 2015.

[272] E. Mattei, D. Lozzi, A. D. Matteo, C. Manes, F. Mignosi, M. Polsinelli, and G. Placidi, "Analysis of the eeg resting-state signals for bci," in *9th Graz Brain-Computer Interface Conference 2024*, pp. 355–359, 2024.

[273] K. S. Kassam, A. R. Markey, V. L. Cherkassky, G. Loewenstein, and M. A. Just, "Identifying emotions on the basis of neural activation," *PloS one*, vol. 8, no. 6, p. e66032, 2013.

[274] E. P. Torres, E. A. Torres, M. Hernández-Álvarez, and S. G. Yoo, "Eeg-based bci emotion recognition: A survey," *Sensors*, vol. 20, no. 18, p. 5083, 2020.

[275] G. Placidi, D. Avola, A. Petracca, F. Sgallari, and M. Spezialetti, "Basis for the implementation of an eeg-based single-trial binary brain computer interface through the disgust produced by remembering unpleasant odors," *Neurocomputing*, vol. 160, pp. 308–318, 2015.

[276] G. Placidi, A. Petracca, M. Spezialetti, and D. Iacoviello, "A modular framework for eeg web based binary brain computer interfaces to recover communication abilities in impaired people," *Journal of medical systems*, vol. 40, no. 1, pp. 1–14, 2016.

[277] J. M. Garcia-Garcia, V. M. Penichet, and M. D. Lozano, "Emotion detection: a technology review," in *Proceedings of the XVIII international conference on human computer interaction*, pp. 1–8, 2017.

[278] M. Spezialetti, L. Cinque, J. M. R. Tavares, and G. Placidi, "Towards eeg-based bci driven by emotions for addressing bci-illiteracy: a meta-analytic review," *Behaviour & Information Technology*, vol. 37, no. 8, pp. 855–871, 2018.

[279] T. Fujimura, Y.-T. Matsuda, K. Katahira, M. Okada, and K. Okanoya, "Categorical and dimensional perceptions in decoding emotional facial expressions," *Cognition & emotion*, vol. 26, no. 4, pp. 587–601, 2012.

[280] Y. Wang, W. Song, W. Tao, A. Liotta, D. Yang, X. Li, S. Gao, Y. Sun, W. Ge, W. Zhang, *et al.*, "A systematic review on affective computing: Emotion models, databases, and recent advances," *Information Fusion*, vol. 83, pp. 19–52, 2022.

[281] H. Liu, Y. Zhang, Y. Li, and X. Kong, "Review on emotion recognition based on electroencephalography," *Frontiers in Computational Neuroscience*, vol. 15, p. 84, 2021.

[282] G. Placidi, M. Polsinelli, M. Spezialetti, and L. Cinque, "Bci driven by self-induced emotions: a multi-class study," in *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pp. 1–6, IEEE, 2018.

[283] S. M. Alarcao and M. J. Fonseca, "Emotions recognition using eeg signals: A survey," *IEEE Transactions on Affective Computing*, vol. 10, no. 3, pp. 374–393, 2017.

[284] D. Iacoviello, A. Petracca, M. Spezialetti, and G. Placidi, "A classification algorithm for electroencephalography signals by self-induced emotional stimuli," *IEEE transactions on cybernetics*, vol. 46, no. 12, pp. 3171–3180, 2015.

[285] A. Dzedzickis, A. Kaklauskas, and V. Bucinskas, "Human emotion recognition: Review of sensors and methods," *Sensors*, vol. 20, no. 3, p. 592, 2020.

[286] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.

[287] C. Neuper, G. R. Müller-Putz, R. Scherer, and G. Pfurtscheller, "Motor imagery and eeg-based control of spelling devices and neuroprostheses," *Progress in brain research*, vol. 159, pp. 393–409, 2006.

[288] A. Kübler, A. Furdea, S. Halder, E. M. Hammer, F. Nijboer, and B. Kotchoubey, "A brain–computer interface controlled auditory event-related potential (p300) spelling system for locked-in patients," *Annals of the New York Academy of Sciences*, vol. 1157, no. 1, pp. 90–100, 2009.

[289] G. R. Muller-Putz, R. Scherer, C. Neuper, and G. Pfurtscheller, "Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces?," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 14, no. 1, pp. 30–37, 2006.

[290] G. Placidi, L. Cinque, P. Di Giamberardino, D. Iacoviello, and M. Spezialetti, "An affective bci driven by self-induced emotions for people with severe neurological disorders," in *New Trends in Image Analysis and Processing–ICIAP 2017: ICIAP International Workshops, WBICV, SSPandBE, 3AS, RGBD, NIVAR, IWBAAS, and MADiMa 2017, Catania, Italy, September 11-15, 2017, Revised Selected Papers 19*, pp. 155–162, Springer, 2017.

[291] G. Garcia-Molina, T. Tsoneva, and A. Nijholt, "Emotional brain–computer interfaces," *International journal of autonomous and adaptive communications systems*, vol. 6, no. 1, pp. 9–25, 2013.

[292] D. Nie, X.-W. Wang, L.-C. Shi, and B.-L. Lu, "Eeg-based emotion recognition during watching movies," in *2011 5th international IEEE/EMBS conference on neural engineering*, pp. 667–670, IEEE, 2011.

[293] D. Iacoviello, A. Petracca, M. Spezialetti, and G. Placidi, "A real-time classification algorithm for eeg-based bci driven by self-induced emotions," *Computer methods and programs in biomedicine*, vol. 122, no. 3, pp. 293–303, 2015.

[294] D. Iacoviello, A. Petracca, M. Spezialetti, and G. Placidi, "A classification algorithm for electroencephalography signals by self-induced emotional stimuli," *IEEE transactions on cybernetics*, vol. 46, no. 12, pp. 3171–3180, 2015.

[295] F. Pistoia, A. Carolei, D. Iacoviello, A. Petracca, S. Sacco, M. Sarà, M. Spezialetti, and G. Placidi, "Eeg-detected olfactory imagery to reveal covert consciousness in minimally conscious state," *Brain injury*, vol. 29, no. 13-14, pp. 1729–1735, 2015.

[296] R. W. Picard and J. Klein, "Computers that recognise and respond to user emotion: theoretical and practical implications," *Interacting with computers*, vol. 14, no. 2, pp. 141–169, 2002.

[297] H. Kober, L. F. Barrett, J. Joseph, E. Bliss-Moreau, K. Lindquist, and T. D. Wager, "Functional grouping and cortical–subcortical interactions in emotion: a meta-analysis of neuroimaging studies," *Neuroimage*, vol. 42, no. 2, pp. 998–1031, 2008.

[298] K. A. Lindquist, T. D. Wager, H. Kober, E. Bliss-Moreau, and L. F. Barrett, "The brain basis of emotion: a meta-analytic review," *Behavioral and brain sciences*, vol. 35, no. 3, pp. 121–143, 2012.

[299] S. M. Alarcao and M. J. Fonseca, "Emotions recognition using eeg signals: A survey," *IEEE transactions on affective computing*, vol. 10, no. 3, pp. 374–393, 2017.

[300] K. Ochsner and J. Gross, "The cognitive control of emotion, trends in cognitive sciences. 2005; 9 (5): 242-249."

[301] G. Placidi, D. Avola, A. Petracca, F. Sgallari, and M. Spezialetti, "Basis for the implementation of an eeg-based single-trial binary brain computer interface through the disgust produced by remembering unpleasant odors," *Neurocomputing*, vol. 160, p. 308–318, July 2015.

[302] J. A. Russell, "A circumplex model of affect.," *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[303] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *Journal of research in Personality*, vol. 11, no. 3, pp. 273–294, 1977.

[304] J. R. Fontaine, K. R. Scherer, E. B. Roesch, and P. C. Ellsworth, "The world of emotions is not two-dimensional," *Psychological science*, vol. 18, no. 12, pp. 1050–1057, 2007.

[305] K. Kamble and J. Sengupta, "A comprehensive survey on emotion recognition based on electroencephalograph (eeg) signals," *Multimedia Tools and Applications*, pp. 1–36, 2023.

[306] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 42–55, 2011.

[307] S. Katsigiannis and N. Ramzan, "Dreamer: A database for emotion recognition through eeg and ecg signals from wireless low-cost off-the-shelf devices," *IEEE journal of biomedical and health informatics*, vol. 22, no. 1, pp. 98–107, 2017.

[308] J. A. Onton and S. Makeig, "High-frequency broadband modulation of electroencephalographic spectra," *Frontiers in human neuroscience*, vol. 3, p. 560, 2009.

[309] A. D. Bigirimana, N. H. Siddique, and D. Coyle, "Brain-computer interfacing with emotion-inducing imagery: A pilot study," in *the 7th Graz BCI Conference 2017*, TU Graz, 2017.

[310] A. Marchewka, Ł. Żurawski, K. Jednoróg, and A. Grabowska, "The nencki affective picture system (naps): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database," *Behavior research methods*, vol. 46, pp. 596–610, 2014.

[311] P. Lang and M. M. Bradley, "The international affective picture system (iaps) in the study of emotion and attention," *Handbook of emotion elicitation and assessment*, vol. 29, pp. 70–73, 2007.

[312] P. Sarma and S. Barma, "Review on stimuli presentation for affect analysis based on eeg," *IEEE Access*, vol. 8, pp. 51991–52009, 2020.

[313] M. Soleymani, J. J. Kierkels, G. Chanel, and T. Pun, "A bayesian framework for video affective representation," in *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pp. 1–7, IEEE, 2009.

[314] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[315] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld, "Learning realistic human actions from movies," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, IEEE, 2008.

[316] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks," *IEEE Transactions on autonomous mental development*, vol. 7, no. 3, pp. 162–175, 2015.

[317] S. B. Eysenck, H. J. Eysenck, and P. Barrett, "A revised version of the psychoticism scale," *Personality and individual differences*, vol. 6, no. 1, pp. 21–29, 1985.

[318] P. Philippot, "Inducing and assessing differentiated emotion-feeling states in the laboratory," *Cognition and emotion*, vol. 7, no. 2, pp. 171–193, 1993.

[319] R. Subramanian, J. Wache, M. K. Abadi, R. L. Vieriu, S. Winkler, and N. Sebe, "Ascertain: Emotion and personality recognition using commercial sensors," *IEEE Transactions on Affective Computing*, vol. 9, no. 2, pp. 147–160, 2016.

[320] M. K. Abadi, R. Subramanian, S. M. Kia, P. Avesani, I. Patras, and N. Sebe, "Decaf: Meg-based multimodal database for decoding affective physiological responses," *IEEE Transactions on Affective Computing*, vol. 6, no. 3, pp. 209–222, 2015.

[321] C. A. Gabert-Quillen, E. E. Bartolini, B. T. Abravanel, and C. A. Sanislow, "Ratings for emotion film clips," *Behavior research methods*, vol. 47, pp. 773–787, 2015.

[322] N. A. Badcock, P. Mousikou, Y. Mahajan, P. De Lissa, J. Thie, and G. McArthur, "Validation of the emotiv epoc® eeg gaming system for measuring research quality auditory erps," *PeerJ*, vol. 1, p. e38, 2013.

[323] C. Kothe, "The artifact subspace reconstruction method," *Accessed: Jul*, vol. 17, p. 2017, 2013.

[324] W. Zheng, W. Liu, Y. Lu, B. Lu, and A. Cichocki, "Emotionmeter: A multimodal framework for recognizing human emotions," *IEEE Transactions on Cybernetics*, pp. 1–13, 2018.

[325] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: the panas scales.," *Journal of personality and social psychology*, vol. 54, no. 6, p. 1063, 1988.

[326] J. A. Miranda-Correa, M. K. Abadi, N. Sebe, and I. Patras, "Amigos: A dataset for affect, personality and mood research on individuals and groups," *IEEE transactions on affective computing*, vol. 12, no. 2, pp. 479–493, 2018.

[327] G. Gómez-Herrero, K. Rutanen, and K. Egiazarian, "Blind source separation by entropy rate minimization," *IEEE signal processing letters*, vol. 17, no. 2, pp. 153–156, 2009.

[328] T.-H. Li, W. Liu, W.-L. Zheng, and B.-L. Lu, "Classification of five emotions from eeg and eye movement signals: Discrimination ability and stability over time," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 607–610, IEEE, 2019.

[329] S. Sangnark, P. Autthasan, P. Ponglertnapakorn, P. Chalekarn, T. Sudhawiyangkul, M. Trakulruangroj, S. Songsermsawad, R. Assabumrungrat, S. Amplod, K. Ounjai, *et al.*, "Revealing preference in popular music through familiarity and brain response," *IEEE Sensors Journal*, vol. 21, no. 13, pp. 14931–14940, 2021.

[330] X. Wang, C. Yan, H. Zheng, Y. Li, and C. Wang, "Design and optimization of actuator for multi-joint soft rehabilitation glove," *Industrial Robot the International Journal of Robotics Research and Application*, vol. 48, no. 6, pp. 877–890, 2021.

[331] M. M. Rahman, A. K. Sarkar, M. A. Hossain, M. S. Hossain, M. R. Islam, M. B. Hossain, J. M. Quinn, and M. A. Moni, "Recognition of human emotions using eeg signals: A review," *Computers in Biology and Medicine*, vol. 136, p. 104696, Sept. 2021.

[332] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 15, p. 056013, July 2018.

[333] Y. Wang, S. Qiu, J. Li, X. Ma, Z. Liang, H. Li, and H. He, "Eeg-based emotion recognition with similarity learning network," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, July 2019.

[334] T. Song, W. Zheng, P. Song, and Z. Cui, "Eeg emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, vol. 11, p. 532–541, July 2020.

[335] T. Song, S. Liu, W. Zheng, Y. Zong, and Z. Cui, "Instance-adaptive graph for eeg emotion recognition," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, p. 2701–2708, Apr. 2020.

[336] Y. Li, W. Zheng, Y. Zong, Z. Cui, T. Zhang, and X. Zhou, "A bi-hemisphere domain adversarial neural network model for eeg emotion recognition," *IEEE Transactions on Affective Computing*, vol. 12, no. 2, pp. 494–504, 2018.

[337] Y. Li, L. Wang, W. Zheng, Y. Zong, L. Qi, Z. Cui, T. Zhang, and T. Song, "A novel bi-hemispheric discrepancy model for eeg emotion recognition," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 2, pp. 354–367, 2020.

[338] B.-Q. Ma, H. Li, W.-L. Zheng, and B.-L. Lu, "Reducing the subject variability of eeg signals with adversarial domain generalization," in *Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12–15, 2019, Proceedings, Part I 26*, pp. 30–42, Springer, 2019.

[339] J. Li, S. Qiu, C. Du, Y. Wang, and H. He, "Domain adaptation for eeg emotion recognition based on latent representation similarity," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 12, no. 2, pp. 344–353, 2019.

[340] Z. He, Y. Zhong, and J. Pan, "An adversarial discriminative temporal convolutional network for eeg-based cross-domain emotion recognition," *Computers in Biology and Medicine*, vol. 141, p. 105048, Feb. 2022.

[341] P. Zhong, D. Wang, and C. Miao, "Eeg-based emotion recognition using regularized graph neural networks," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1290–1301, 2020.

[342] Y. Cimtay and E. Ekmekcioglu, "Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset eeg emotion recognition," *Sensors*, vol. 20, no. 7, p. 2034, 2020.

[343] H. Chen, S. Sun, J. Li, R. Yu, N. Li, X. Li, and B. Hu, "Personal-zscore: Eliminating individual difference for eeg-based cross-subject emotion recognition," *IEEE Transactions on Affective Computing*, vol. 14, no. 3, pp. 2077–2088, 2021.

[344] J. Cao, X. He, C. Yang, and Z. Wang, "Multi-source and multi-representation adaptation for cross-domain electroencephalography emotion recognition," *Frontiers in Psychology*, vol. 12, p. 809459, 2022.

[345] H. Cui, A. Liu, X. Zhang, X. Chen, J. Liu, and X. Chen, "Eeg-based subject-independent emotion recognition using gated recurrent unit and minimum class confusion," *IEEE Transactions on Affective Computing*, 2022.

[346] Z. Li, E. Zhu, M. Jin, C. Fan, H. He, T. Cai, and J. Li, "Dynamic domain adaptation for class-aware cross-subject and cross-session eeg emotion recognition," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 12, pp. 5964–5973, 2022.

[347] G. Zhang, V. Davoodnia, and A. Etemad, "Parse: Pairwise alignment of representations in semi-supervised eeg learning for emotion recognition," *IEEE Transactions on Affective Computing*, vol. 13, no. 4, pp. 2185–2200, 2022.

[348] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks," *IEEE Transactions on autonomous mental development*, vol. 7, no. 3, pp. 162–175, 2015.

[349] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "Emotionmeter: A multimodal framework for recognizing human emotions," *IEEE transactions on cybernetics*, vol. 49, no. 3, pp. 1110–1122, 2018.

[350] S.-E. Moon, S. Jang, and J.-S. Lee, "Convolutional neural network approach for eeg-based emotion recognition using brain connectivity and its spatial information," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2556–2560, IEEE, 2018.

[351] S. Sun, X. Li, J. Zhu, Y. Wang, R. La, X. Zhang, L. Wei, and B. Hu, "Graph theory analysis of functional connectivity in major depression disorder with high-density resting state eeg data," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 3, pp. 429–439, 2019.

[352] Y. Peng, Y. Huang, B. Chen, M. He, L. Jiang, Y. Li, X. Huang, C. Pei, S. Zhang, C. Li, *et al.*, "Electroencephalographic network topologies predict antidepressant responses in patients with major depressive disorder," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2577–2588, 2022.

[353] A. Maghsoudi and A. Shalbaf, "Hand motor imagery classification using effective connectivity and hierarchical machine learning in eeg signals," *Journal of Biomedical Physics & Engineering*, vol. 12, no. 2, p. 161, 2022.

[354] L. E. Ismail and W. Karwowski, "Applications of eeg indices for the quantification of human cognitive performance: A systematic review and bibliometric analysis," *Plos one*, vol. 15, no. 12, p. e0242857, 2020.

[355] W. Kong, L. Wang, S. Xu, F. Babiloni, and H. Chen, "Eeg fingerprints: phase synchronization of eeg signals as biomarker for subject identification," *IEEE Access*, vol. 7, pp. 121165–121173, 2019.

[356] M. A. Hendrawan, U. D. Rosiani, and A. D. W. Sumari, "Single channel electroencephalogram (eeg) based biometric system," in *2022 IEEE 8th Information Technology International Seminar (ITIS)*, pp. 307–311, IEEE, 2022.

[357] W. Alsumari, M. Hussain, L. Alshehri, and H. A. Aboalsamh, "Eeg-based person identification and authentication using deep convolutional neural network," *Axioms*, vol. 12, no. 1, p. 74, 2023.

[358] J. A. Onton and S. Makeig, "High-frequency broadband modulation of electroencephalographic spectra," *Frontiers in human neuroscience*, p. 61, 2009.

[359] Y. Li, Y. Wang, B. Zhang, Y. Wang, and X. Zhou, "Electrophysiological responses to expectancy violations in semantic and gambling tasks: a comparison of different eeg reference approaches," *Frontiers in neuroscience*, vol. 12, p. 169, 2018.

[360] D. Karamacoska, R. J. Barry, and G. Z. Steiner, "Resting state intrinsic eeg impacts on go stimulus-response processes," *Psychophysiology*, vol. 54, no. 6, pp. 894–903, 2017.

[361] K. A. Bujarski, Y. Song, T. Xie, Z. Leeds, S. I. Kolankiewicz, G. H. Wozniak, S. Guillory, J. P. Aronson, L. Chang, and B. C. Jobst, "Modulation of emotion perception via amygdala stimulation in humans," *Frontiers in Neuroscience*, vol. 15, p. 795318, 2022.

[362] R. B. Weldon, K. Y. Behrens, D. K. Jones, E. A. Drake, and J. R. Fragetta, "Anxiety and avoidance in close relationships are associated with responses to negative and positive stimuli," *Psychological Reports*, vol. 126, no. 6, pp. 2821–2833, 2023.

[363] G. E. Barnacle, D. Tsivilis, A. Schaefer, and D. Talmi, "Local context influences memory for emotional stimuli but not electrophysiological markers of emotion-dependent attention," *Psychophysiology*, vol. 55, no. 4, p. e13014, 2018.

[364] R. K. Chikara, O. Komarov, and L.-W. Ko, "Neural signature of event-related n200 and p300 modulation in parietal lobe during human response inhibition," *International Journal of Computational Biology and Drug Design*, vol. 11, no. 1-2, pp. 171–182, 2018.

[365] J. H. Ford, J. A. Morris, and E. A. Kensinger, "Effects of emotion and emotional valence on the neural correlates of episodic memory search and elaboration," *Journal of Cognitive Neuroscience*, vol. 26, no. 4, pp. 825–839, 2014.

[366] M. D. Eddy, T. T. Brunyé, S. Tower-Richardi, C. R. Mahoney, and H. A. Taylor, "The effect of a brief mindfulness induction on processing of emotional images: an erp study," *Frontiers in Psychology*, vol. 6, p. 1391, 2015.

[367] L. Dong, X. Liu, L. Zhao, Y. Lai, D. Gong, T. Liu, and D. Yao, "A comparative study of different eeg reference choices for event-related potentials extracted by independent component analysis," *Frontiers in neuroscience*, vol. 13, p. 1068, 2019.

[368] M. E. Cano, Q. A. Class, and J. Polich, "Affective valence, stimulus attributes, and p300: color vs. black/white and normal vs. scrambled images," *International Journal of Psychophysiology*, vol. 71, no. 1, pp. 17–24, 2009.

[369] L. A. Barradas-Chacón, C. Brunner, and S. C. Wriessnegger, "Stylized faces enhance erp features used for the detection of emotional responses," *Frontiers in Human Neuroscience*, vol. 17, p. 1160800, 2023.

[370] S. Zentner, A. Barradas Chacon, and S. C. Wriessnegger, "The impact of light conditions on neural affect classification: A deep learning approach," *Machine Learning and Knowledge Extraction*, vol. 6, no. 1, pp. 199–214, 2024.

[371] Z. Zhang, Y. Peng, and T. Chen, "Om chanting modulates the processing of negative stimuli: Behavioral and electrophysiological evidence," *Frontiers in Psychology*, vol. 13, p. 943243, 2022.

[372] K. W. Brown, R. J. Goodman, and M. Inzlicht, "Dispositional mindfulness and the attenuation of neural responses to emotional stimuli," *Social cognitive and affective neuroscience*, vol. 8, no. 1, pp. 93–99, 2013.

[373] G. Hajcak, A. MacNamara, and D. M. Olvet, "Event-related potentials, emotion, and emotion regulation: an integrative review," *Developmental neuropsychology*, vol. 35, no. 2, pp. 129–155, 2010.

[374] G. Vecchiato, J. Toppi, L. Astolfi, F. De Vico Fallani, F. Cincotti, D. Mattia, F. Bez, and F. Babiloni, "Spectral eeg frontal asymmetries correlate with the experienced pleasantness of tv commercial advertisements," *Medical & biological engineering & computing*, vol. 49, pp. 579–583, 2011.

[375] A. E. Rehmert and M. A. Kisley, "Can older adults resist the positivity effect in neural responding? the impact of verbal framing on event-related brain potentials elicited by emotional images.," *Emotion*, vol. 13, no. 5, p. 949, 2013.

[376] L. Luther, J. M. Horschig, J. M. van Peer, K. Roelofs, O. Jensen, and M. A. Hagenaars, "Oscillatory brain responses to emotional stimuli are effects related to events rather than states," *Frontiers in Human Neuroscience*, vol. 16, p. 868549, 2023.

[377] M. Balconi, E. Brambilla, and L. Falbo, "Appetitive vs. defensive responses to emotional cues. autonomic measures and brain oscillation modulation," *Brain Research*, vol. 1296, pp. 72–84, 2009.

[378] M. Riegel, Ł. Żurawski, M. Wierzba, A. Moslehi, Ł. Klocek, M. Horvat, A. Grabowska, J. Michałowski, K. Jednoróg, and A. Marchewka, "Characterization of the nencki affective picture system by discrete emotional categories (naps be)," *Behavior research methods*, vol. 48, pp. 600–612, 2016.

[379] J. M. Michałowski, D. Droździel, J. Matuszewski, W. Koziejowski, K. Jednoróg, and A. Marchewka, "The set of fear inducing pictures (sfip): Development and validation in fearful and nonfearful individuals," *Behavior Research Methods*, vol. 49, pp. 1407–1419, 2017.

[380] M. Riegel, A. Moslehi, J. M. Michałowski, Ł. Żurawski, M. Horvat, M. Wypych, K. Jednoróg, and A. Marchewka, "Nencki affective picture system: Cross-cultural study in europe and iran," *Frontiers in Psychology*, vol. 8, p. 274, 2017.

[381] M. Horvat, A. Jović, and K. Burnik, "Investigation of relationships between discrete and dimensional emotion models in affective picture databases using unsupervised machine learning," *Applied Sciences*, vol. 12, no. 15, p. 7864, 2022.

[382] B. B. Briesemeister, L. Kuchinke, and A. M. Jacobs, "Discrete emotion norms for nouns: Berlin affective word list (denn–bawl)," *Behavior research methods*, vol. 43, pp. 441–448, 2011.

[383] J. W. Peirce, "Psychopy—psychophysics software in python," *Journal of neuroscience methods*, vol. 162, no. 1-2, pp. 8–13, 2007.

[384] G. Stoet, "Psytoolkit: A software package for programming psychological experiments using linux," *Behavior research methods*, vol. 42, no. 4, pp. 1096–1104, 2010.

[385] G. Stoet, "Psytoolkit: A novel web-based method for running online questionnaires and reaction-time experiments," *Teaching of Psychology*, vol. 44, no. 1, pp. 24–31, 2017.

[386] S. Matlab, *Matlab*, 2012.

[387] C. Smyth, "The pittsburgh sleep quality index (psqi)," 1999.

[388] C. D. Spielberger, F. Gonzalez-Reigosa, A. Martinez-Urrutia, L. F. Natalicio, and D. S. Natalicio, "The state-trait anxiety inventory," *Revista Interamericana de Psicologia/Interamerican journal of psychology*, vol. 5, no. 3 & 4, 1971.

[389] A. T. Beck, R. A. Steer, G. K. Brown, *et al.*, *Beck depression inventory*. Harcourt Brace Jovanovich New York:, 1987.

[390] D. Leising, T. Grande, and R. Faber, "The toronto alexithymia scale (tas-20): A measure of general psychological distress," *Journal of research in personality*, vol. 43, no. 4, pp. 707–710, 2009.

[391] E. Jorgenson, "Development of the open hemispheric brain dominance scale," 2015.

[392] R. J. Craig, "Millon clinical multiaxial inventory-iii," in *Personality assessment*, pp. 145–178, Routledge, 2011.

[393] N. Palomero-Gallagher and K. Amunts, "A short review on emotion processing: A lateralized network of neuronal networks," *Brain Structure and Function*, vol. 227, no. 2, pp. 673–684, 2022.

[394] P. Ekman *et al.*, "Basic emotions," *Handbook of cognition and emotion*, vol. 98, no. 45-60, p. 16, 1999.

[395] M. Kutas and K. D. Federmeier, "Thirty years and counting: finding meaning in the n400 component of the event-related brain potential (erp)," *Annual review of psychology*, vol. 62, no. 1, pp. 621–647, 2011.

[396] H. U. Amin, A. S. Malik, N. Badruddin, and W.-T. Chooi, "Brain behavior in learning and memory recall process: A high-resolution eeg analysis," in *The 15th International Conference on Biomedical Engineering: ICBME 2013, 4th to 7th December 2013, Singapore*, pp. 683–686, Springer, 2014.

[397] Z. Huang, Y. Ma, R. Wang, W. Li, and Y. Dai, "A model for eeg-based emotion recognition: Cnn-bi-lstm with attention mechanism," *Electronics*, vol. 12, no. 14, p. 3188, 2023.

[398] W. Lu, T.-P. Tan, and H. Ma, "Bi-branch vision transformer network for eeg emotion recognition," *IEEE Access*, vol. 11, pp. 36233–36243, 2023.

[399] R. Xin, F. Miao, P. Cong, F. Zhang, Y. Xin, and X. Feng, "Multiview feature fusion attention convolutional recurrent neural networks for eeg-based emotion recognition," *Journal of Sensors*, vol. 2023, no. 1, p. 9281230, 2023.

[400] D. Watson, "The panas-x: Manual for the positive and negative affect schedule-expanded form," *Unpublished manuscript. University of Iowa, Iowa City*, 1994.

[401] I. Di Pompeo, G. D'Aurizio, C. Burattini, F. Bisegna, and G. Curcio, "Positive mood induction to promote well-being and health: A systematic review from real settings to virtual reality.," *Journal of Environmental Psychology*, p. 102095, 2023.

[402] J. D. Mayer and Y. N. Gaschke, "The brief mood introspection scale (bmis)," 1988.

Part V

Appendix

Appendix A

Sleep Stage Distribution for Chapter 5

	Sleep stage N1	Sleep stage N2	Sleep stage N3	Sleep stage R	Sleep stage W
EEG Fpz-Cz_alpha	False	False	False	False	False
EEG Fpz-Cz_beta	False	False	False	False	False
EEG Fpz-Cz_delta	False	False	False	False	False
EEG Fpz-Cz_gamma	False	False	False	False	False
EEG Fpz-Cz_sigma	False	False	False	False	False
EEG Fpz-Cz_theta	False	False	False	False	False
EEG Pz-Oz_alpha	False	False	False	False	False
EEG Pz-Oz_beta	False	False	False	False	False
EEG Pz-Oz_delta	False	False	False	False	False
EEG Pz-Oz_gamma	False	False	False	False	False
EEG Pz-Oz_sigma	False	False	False	False	False
EEG Pz-Oz_theta	False	False	False	False	False

Table A 1: Shapiro-Wilk Statistic normality test on RBP distribution for each band power and each electrode.

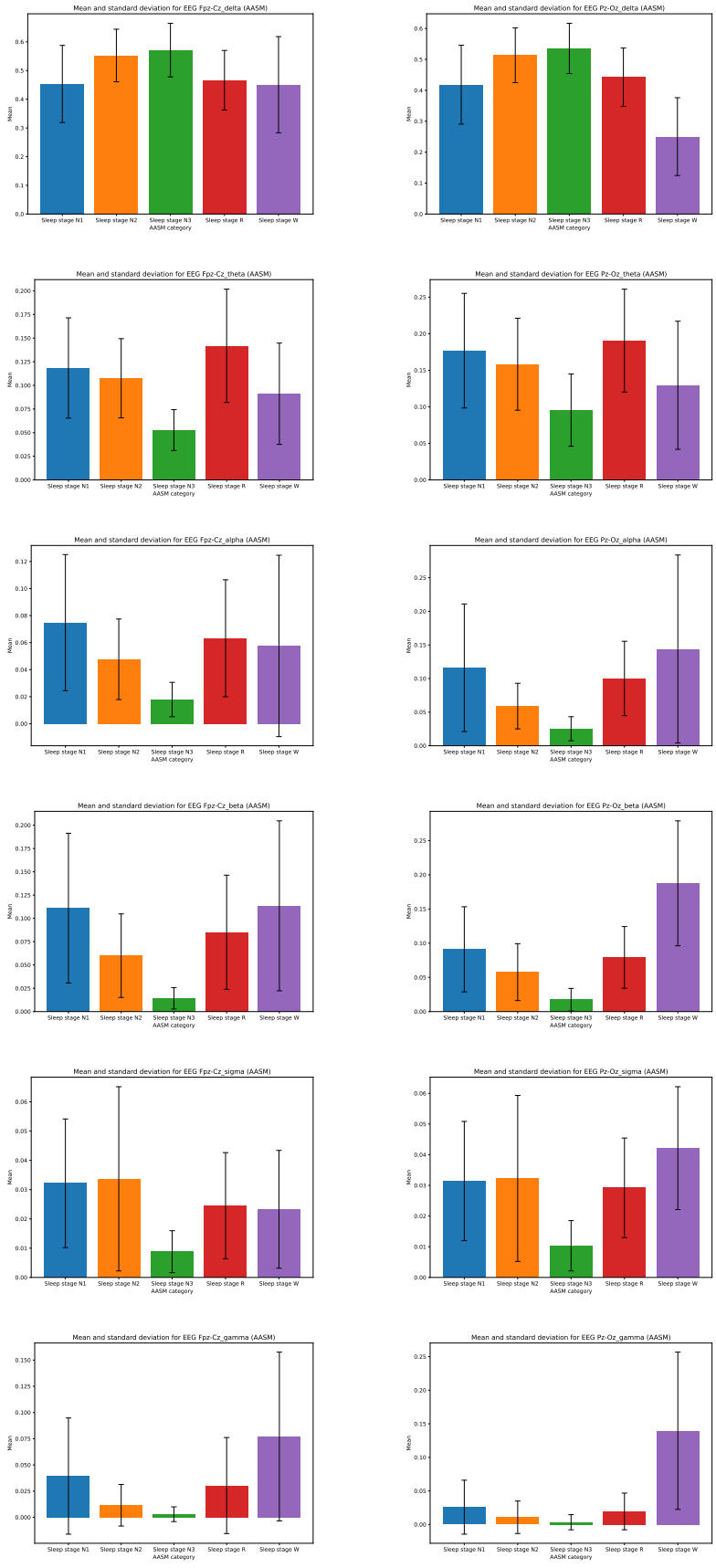


Table A 2: Histograms of the RBP average of sleep stage divided per channels and frequency bands.

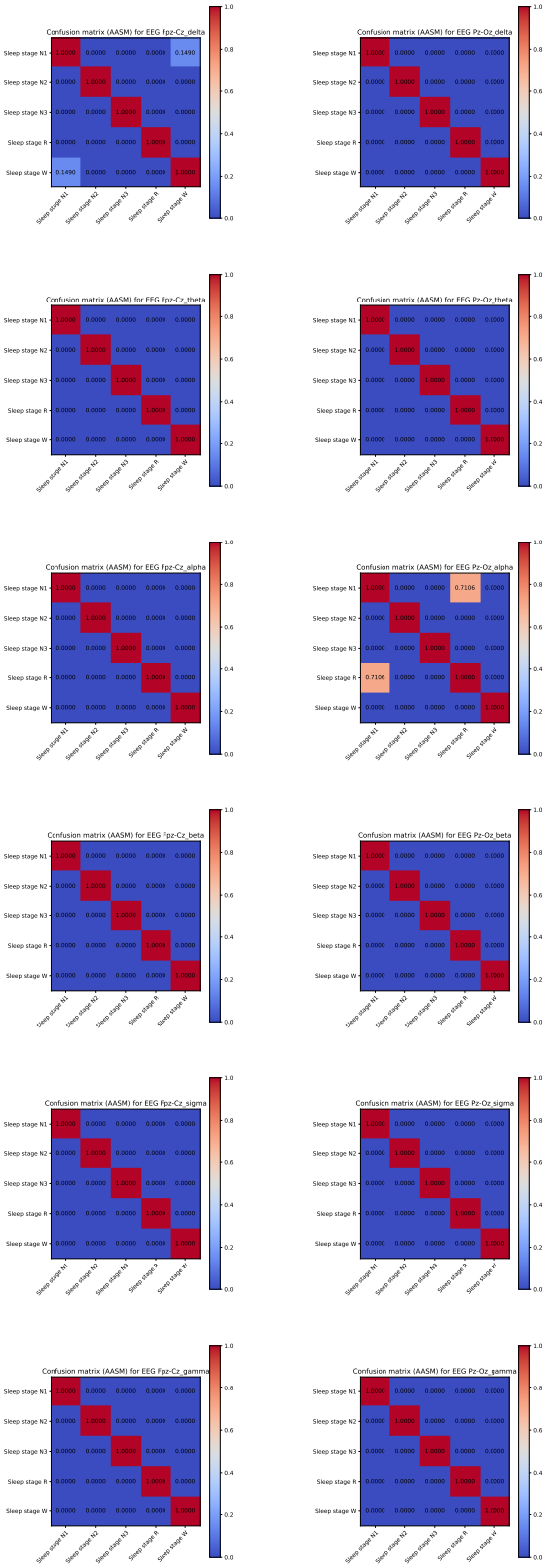


Table A 3: Confusion matrix of the Mann-Whitney test performed on the distribution of the RBP of sleep stage divided per channels and frequency bands.

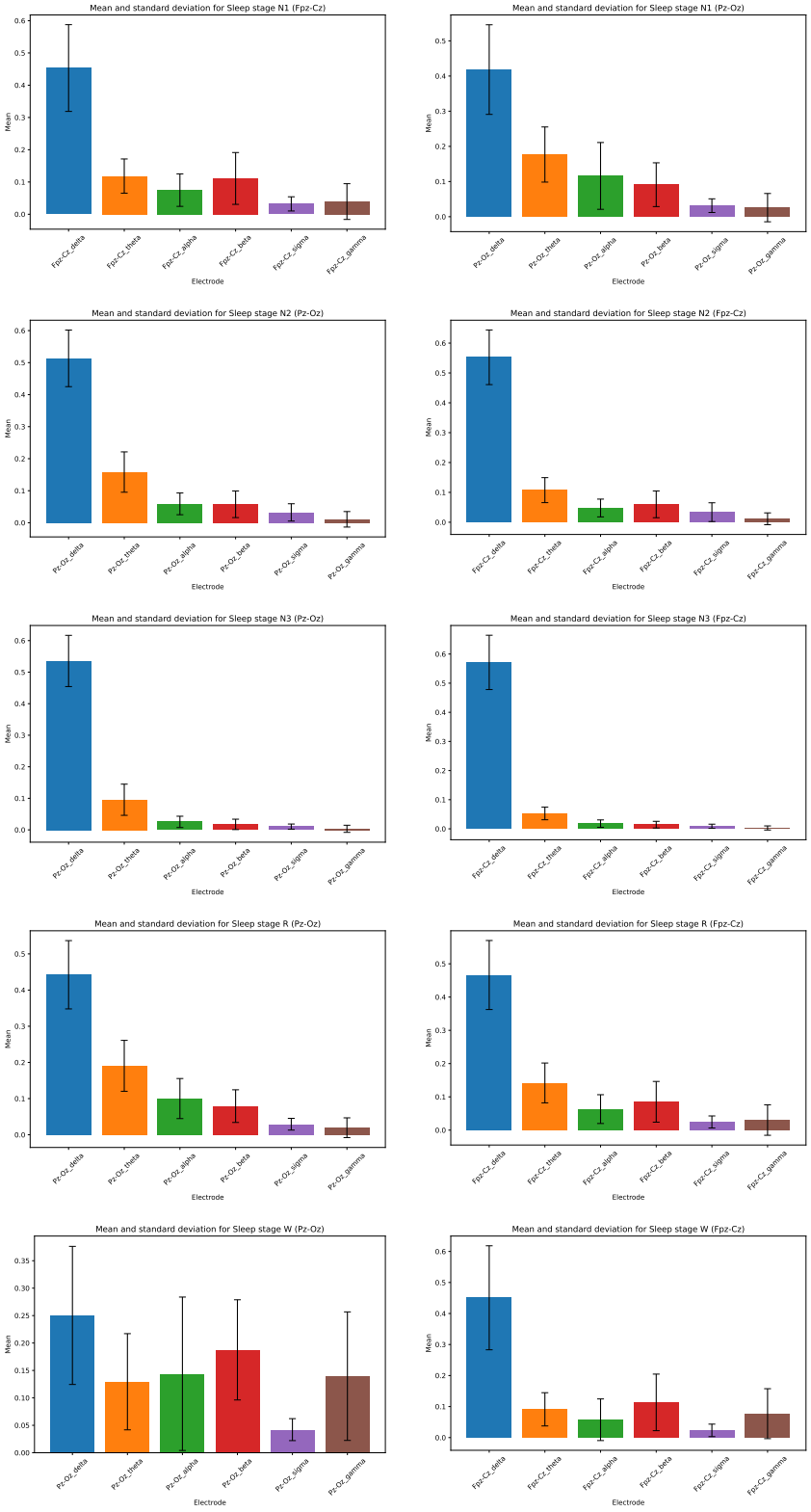


Table A 4: Histograms of the RBP average of sleep stage divided per stage and frequency bands.

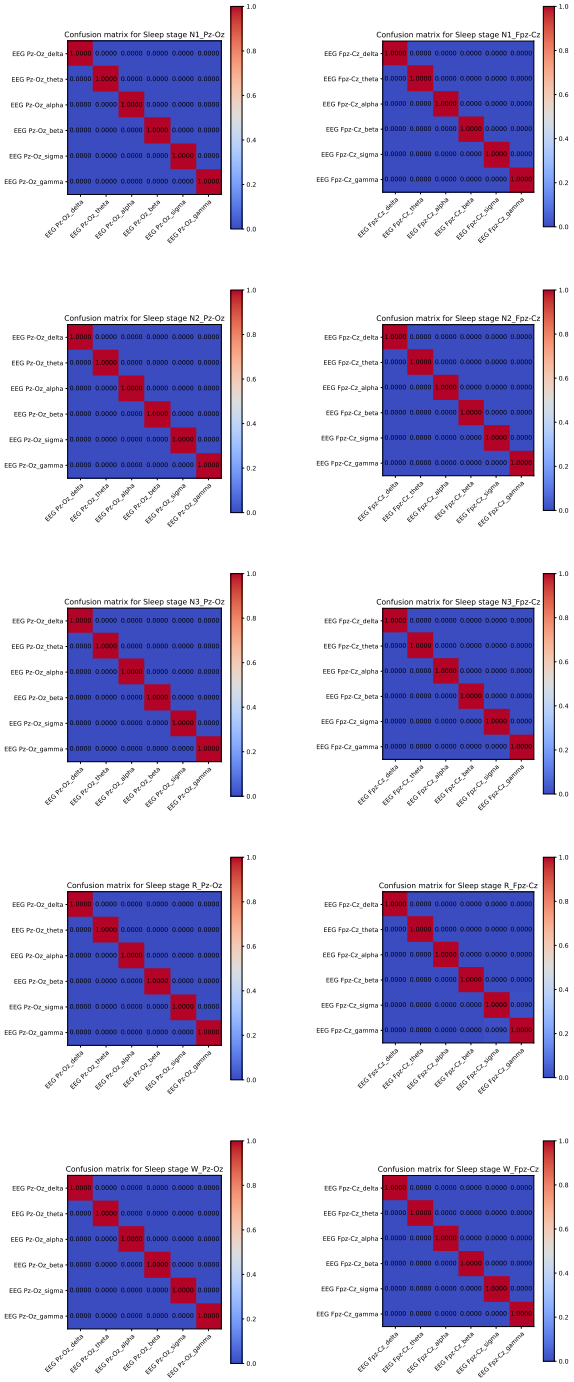


Table A 5: Confusion matrix of the Mann-Whitney test performed on the distribution of the RBP of sleep stage divided per stage and frequency bands.

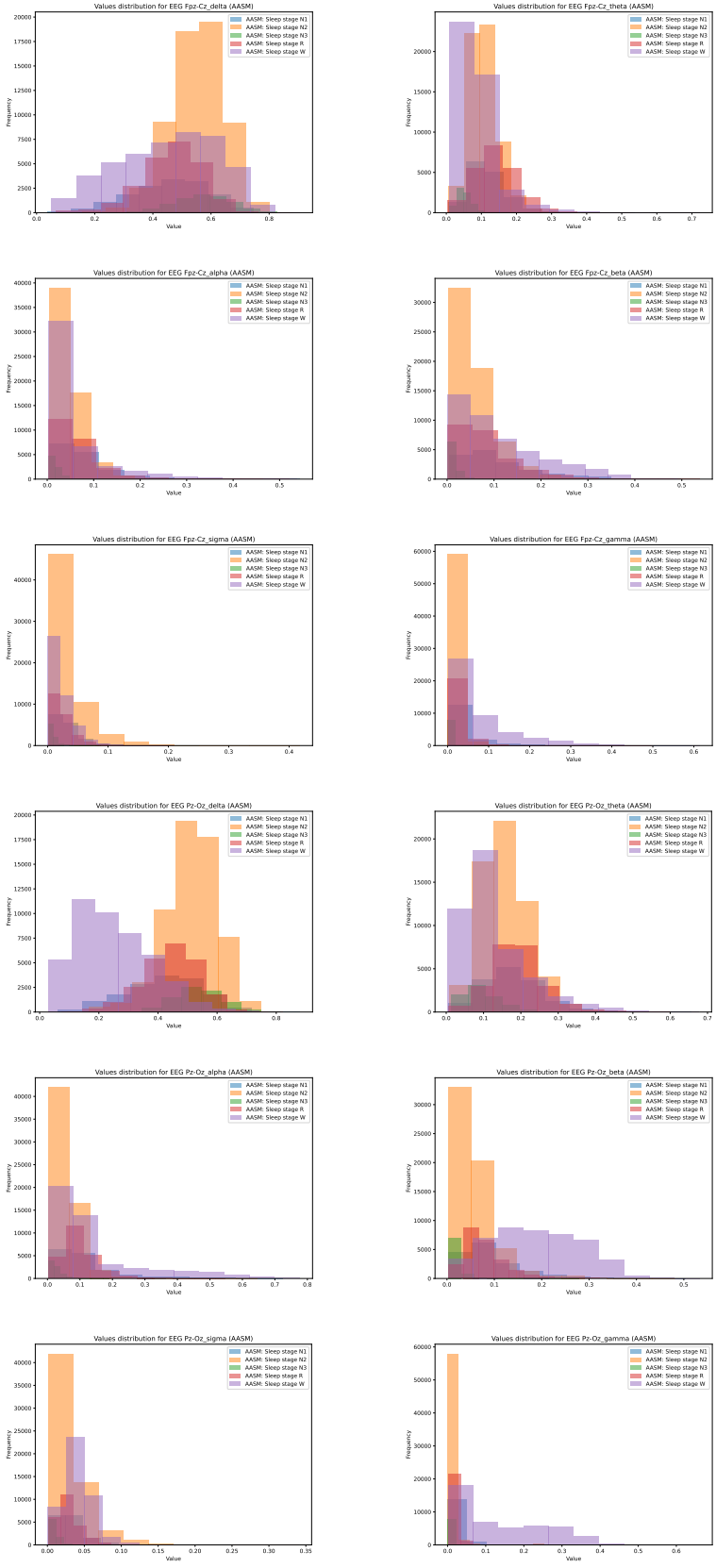


Table A 6: Distribution of RBP of sleep stage divided per channels and frequency bands.

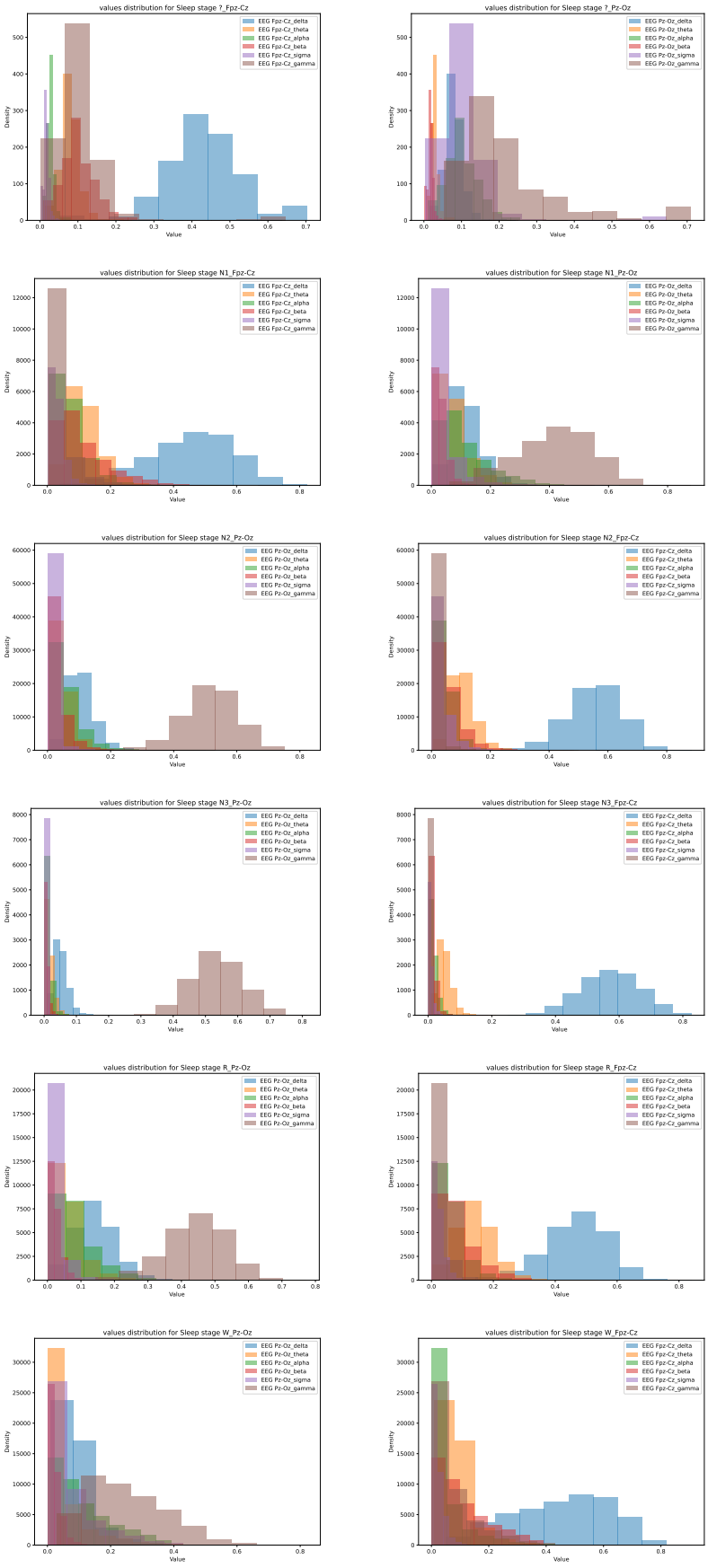


Table A 7: Distribution of RBP of sleep stage divided per channels and sleep stage.

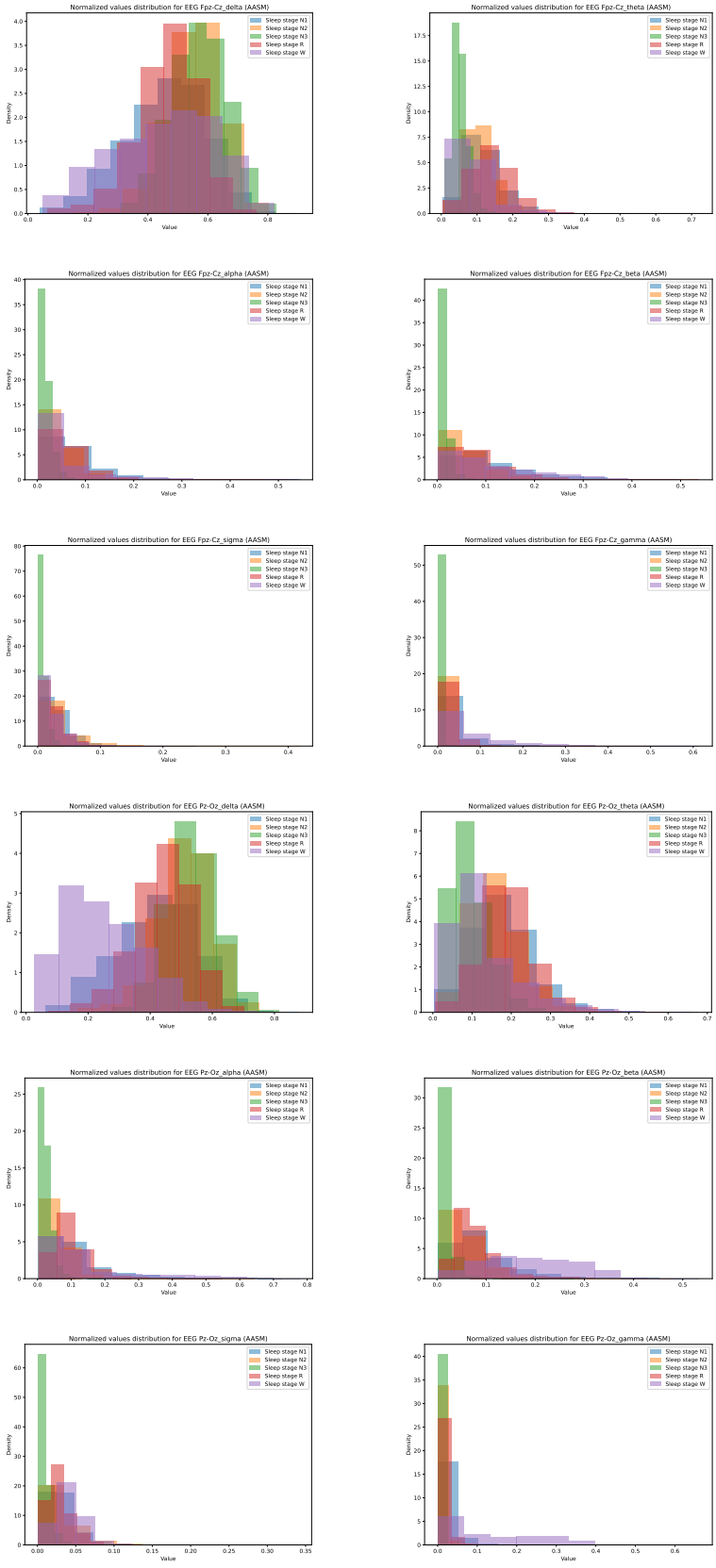


Table A 8: Normalized distribution of RBP of sleep stage divided per channels and frequency bands.

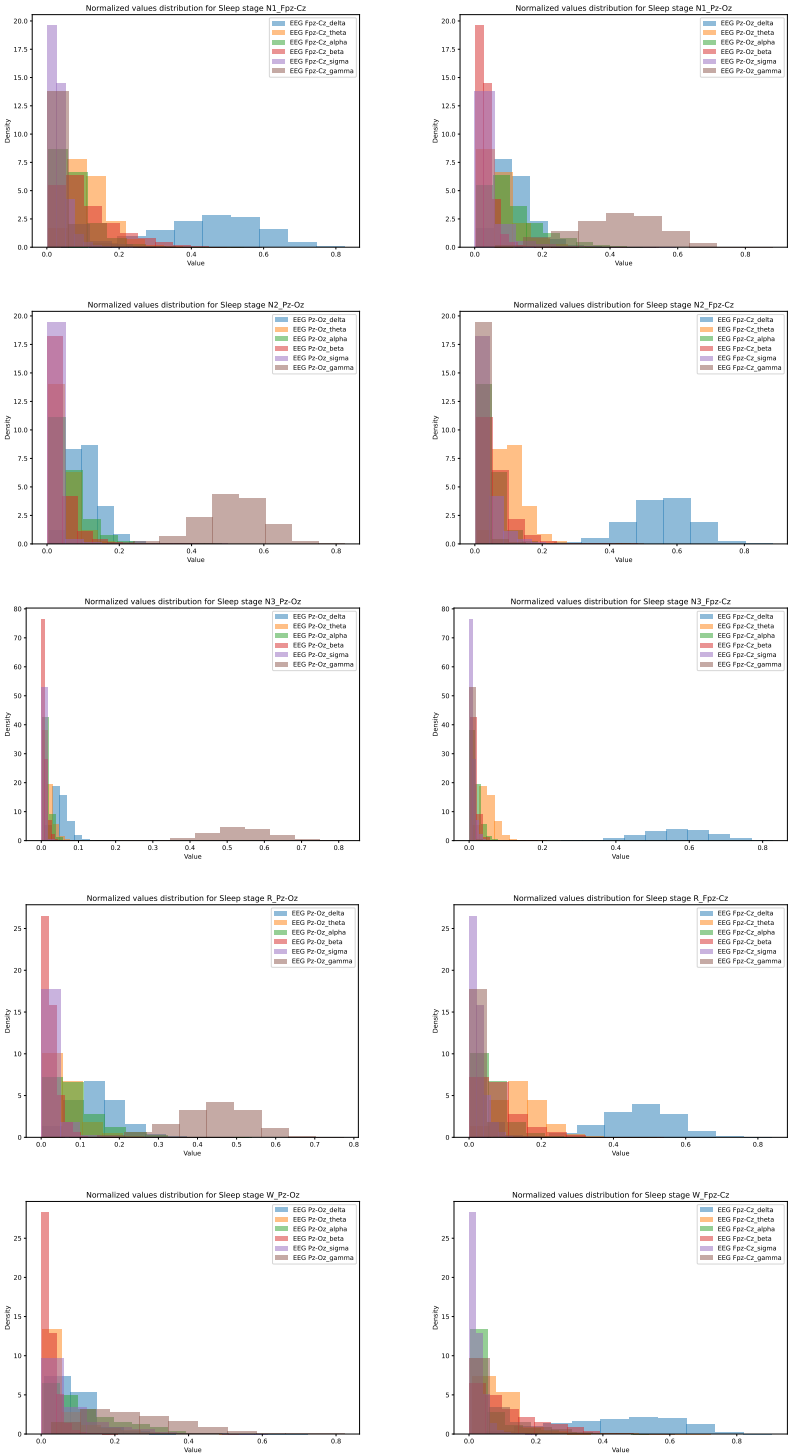


Table A 9: Normalized distribution of RBP of sleep stage divided per channels and sleep stage.

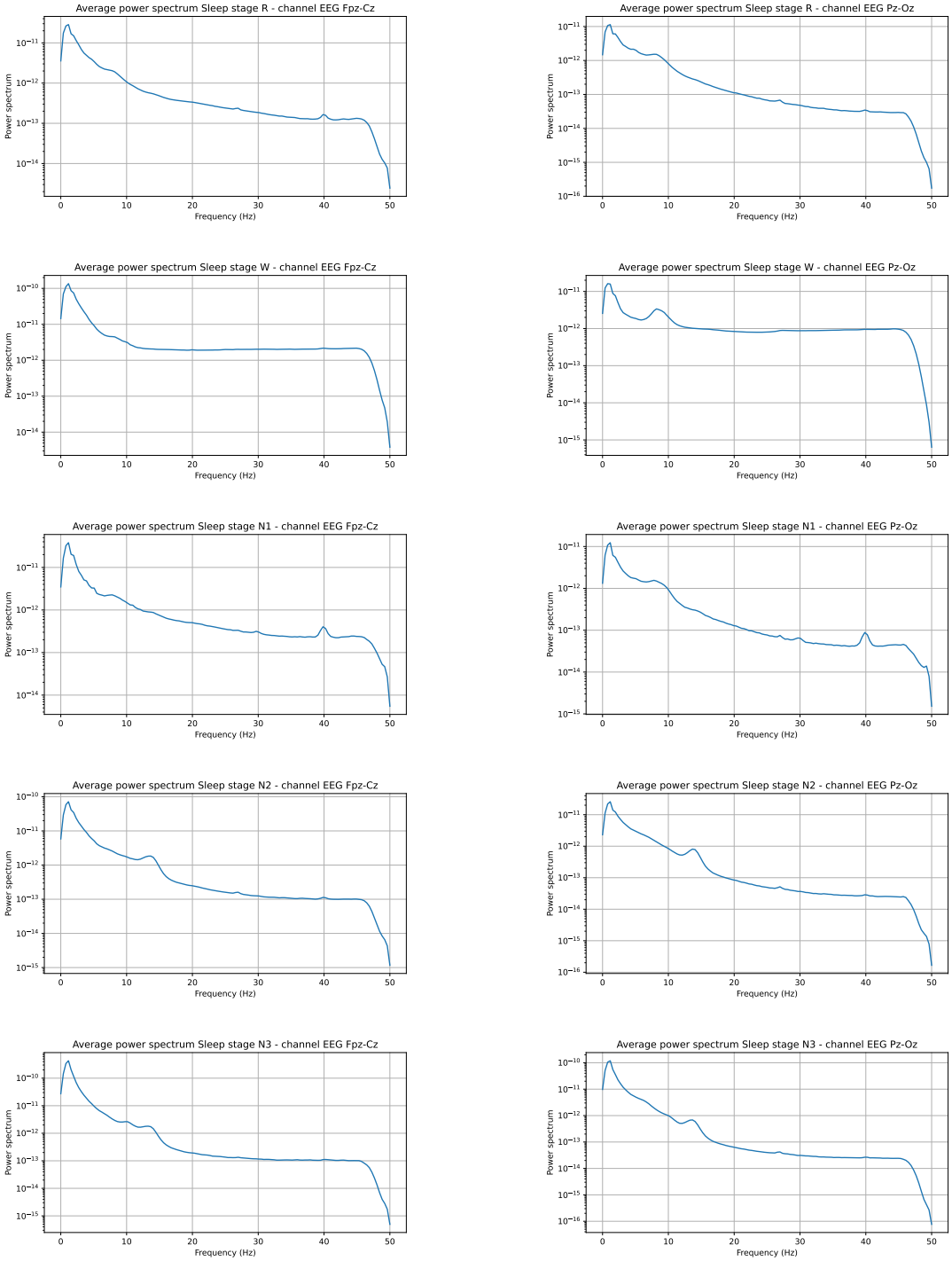


Table A 10: Power spectrum for each channel and sleep stage in AASM classification.

Appendix B

Psychological Questionnaires used in Sections 6.4 and 6.5

Arousal assessment Current value of arousal on a likert scale from 1 to 9.

Valence assessment Current value of valence on a likert scale from 1 to 9.

Dominance assessment Current value of dominance on a likert scale from 1 to 9.

Positive and Negative Affect Schedule-eXpanded (PANAS-X) The PANAS-X [400] scale is a self-administered survey that comprises 60 items on 5-likert point scale, designed to assess both positive and negative emotions. This scale is derived from Positive and Negative Affect Schedule (PANAS) [325]. Participants rate each item on a scale ranging from 1 (*"not experiencing the emotion at all"*) to 5 (*"experiencing the emotion very strongly"*). This is included because is one of the most used scale to assess the affective state in the mood induction experiement [401]. In addition to the two original higher order scales, the PANAS-X measures 11 specific affects:

- Fear (afraid, scared, frightened, nervous, jittery, shaky)
- Hostility (angry, hostile, irritable, scornful, disgusted, loathing)
- Guilt (guilty, ashamed, blameworthy, angry at self, disgusted with self, dissatisfied with self)
- Sadness (sad, blue, downhearted, alone, lonely)
- Joviality (happy, joyful, delighted, cheerful, excited, enthusiastic, lively, energetic)
- Self-Assurance (proud, strong, confident, bold, daring, fearless)
- Attentiveness (alert, attentive, concentrating, determined)
- Shyness (shy, bashful, sheepish, timid)
- Fatigue (sleepy, tired, sluggish, drowsy)
- Serenity (calm, relaxed, at ease)
- Surprise (amazed, surprised, astonished)

Pittsburgh Sleep Quality Index (PSQI) The purpose of the Pittsburgh Sleep Quality Index (PSQI) [387] is to evaluate the general sleep quality in clinical populations. The questionnaire consists of 19 self-reported items, which are divided into seven subcategories: subjective sleep quality, sleep latency, sleep duration, sleep latency habits, habitual sleep efficiency, sleep disturbance, sleep medication use, and diurnal dysfunction. Additionally, there are five extra questions that are answered by the respondent's roommate or bed partner, but these questions are not used for assessment purposes.

State-Trait Anxiety Inventory (STAI) The STAI [388] is a psychological questionnaire that includes 40 items for self-reporting on a 4-point Likert scale. This inventory assesses two categories of anxiety: STAI-Trait (STAI-T) and STAI-State (STAI-S). Increased scores on the inventory indicate a positive correlation with higher levels of anxiety.

Beck Depression Inventory (BDI) The BDI [389] was developed by Aaron T. Beck as a self-report questionnaire consisting of 21 multiple-choice questions. This psychometric test is widely utilized to assess the level of severity of depression and has had a significant impact on the mental health field. Prior to the BDI's development, mental health professionals primarily approached depression from a psychodynamic standpoint, rather than attributing it to the thoughts of the individual.

Toronto Alexithymia Scale (TAS-20) The TAS-20 [390], is a self-administered questionnaire consisting of 20 items. Its purpose is to assess the challenges individuals face in recognizing and articulating their emotions, which is a significant aspect of alexithymia.

Open Hemispheric Brain Dominance Scale (OHBDS) The OHBDS Scale [391] was created by choosing items that had a correlation with five other existing tests that measure left brain and right brain dominance. The process involves using a 24-item inventory, which participants rate based on their level of agreement using a five-point scale. A rating of 1 indicates disagreement, 3 represents "neutrality", and 5 signifies agreement.

Millon Clinical Multiaxial Inventory (MCMI-III) The Millon Clinical Multiaxial Inventory (MCMI-III) [392] is a multidimensional measure of psychopathology, particularly useful for diagnosing personality disorders and dysfunctional relational patterns.

Brief Mood Introspection Scale (BMIS) The BMIS [402] is a public-domain mood scale with 16 adjectives for self-reporting (e.g., Are you "sad"?). It provides scores for overall pleasantness, reflecting the degree to which someone feels positive or negative, and arousal level, indicating how activated or calm they feel, on a 4-likert point scale. Additionally, the BMIS allows for the assessment of specific mood states. These include *pleasant-unpleasant mood*, capturing the range from positive to negative feelings, *arousal-calm mood*, indicating the level of activation versus tranquility, *positive-tired mood*, representing the combination of positive feelings and fatigue, and *negative-relaxed mood*, representing the co-occurrence of negative emotions and relaxation.

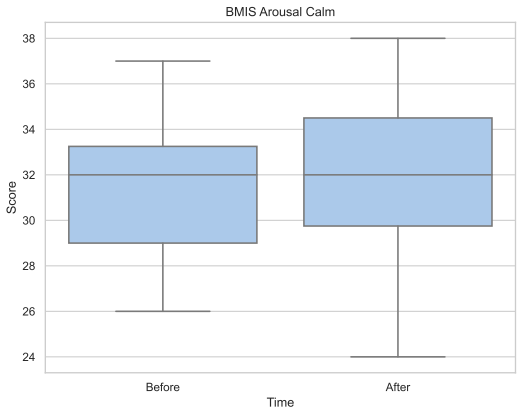
Mood assessment With a sliding bar, participants were asked to give an overall assessment of their mood on a -10 to 10 likert scale.

Self-Assessment Manikin (SAM) A graphical assessment tool that evaluates valence, arousal, and dominance linked to an individual's emotional response to diverse stimuli. It features cartoon figures to portray varying levels of these dimensions, enabling participants to efficiently rate their emotions [314].

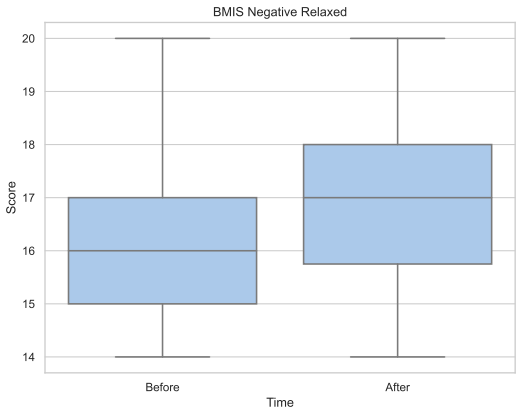
Appendix C

Scoring Psychological Questionnaires for Chapter 6.4

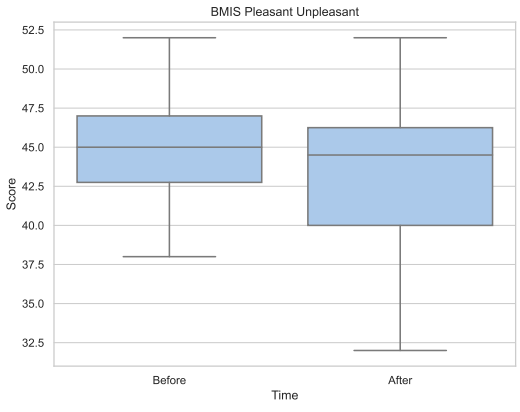
BMIS



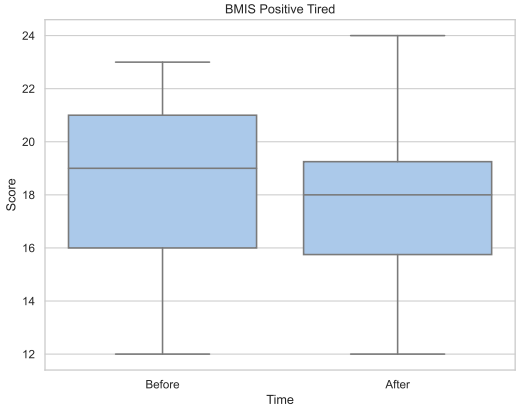
(a)



(b)



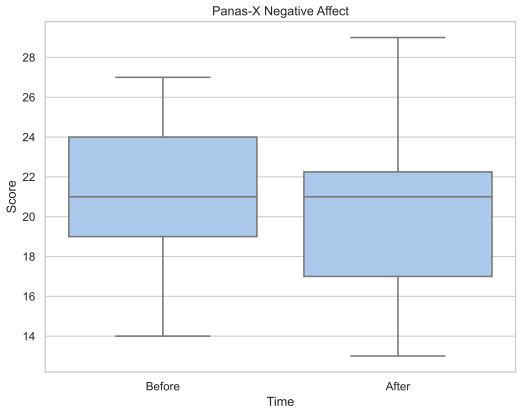
(c)



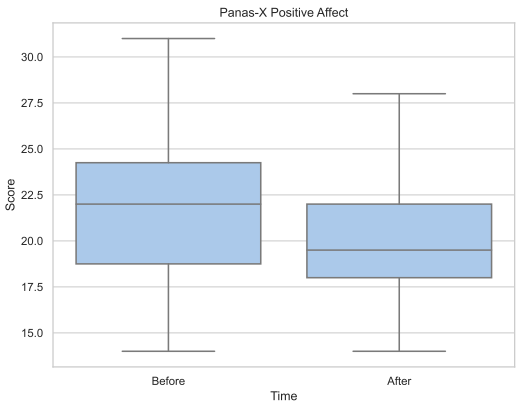
(d)

Table C 1: Brief Mood Introspection Scale (BMIS) results summary before and after the EEG data collection: (a) BMIS Arousal Calm, (b) BMIS Negative Relaxed, (c) BMIS Pleasant Unpleasant and (d) BMIS Positive Tired.

PANAS-X

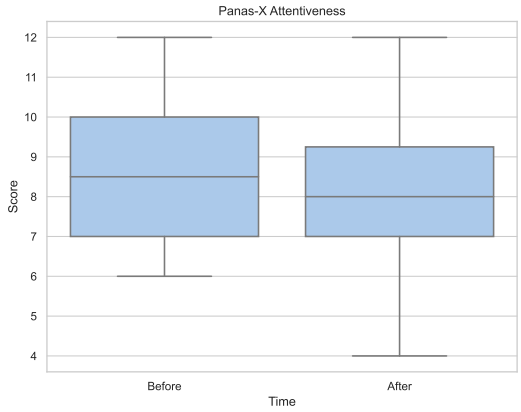


(a)

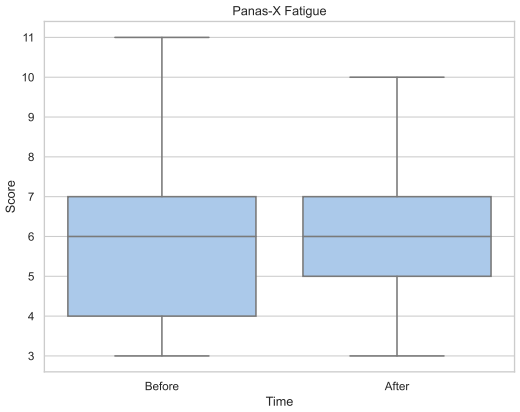


(b)

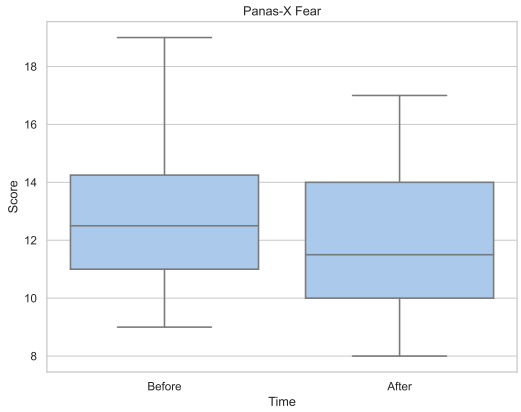
Table C 2: PANAS-X results summary before and after the EEG data collection: (a) Negative Affect and (b) Positive Affect.



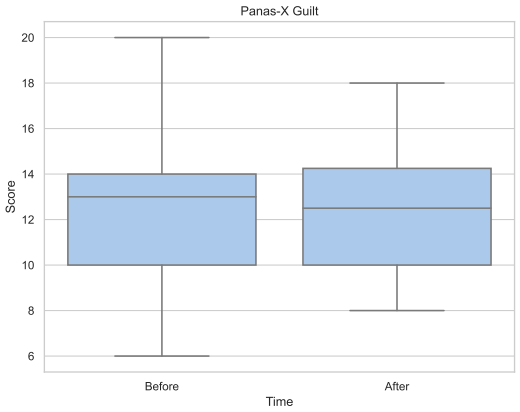
(a) Attentiveness



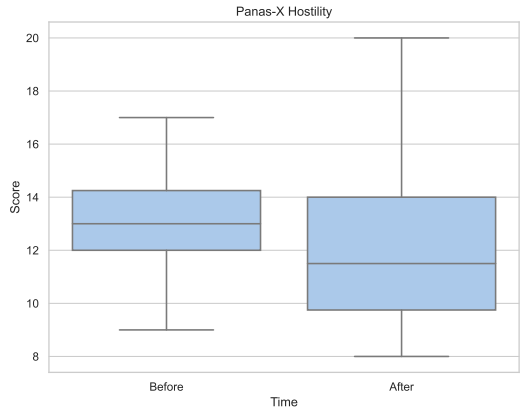
(b) Fatigue



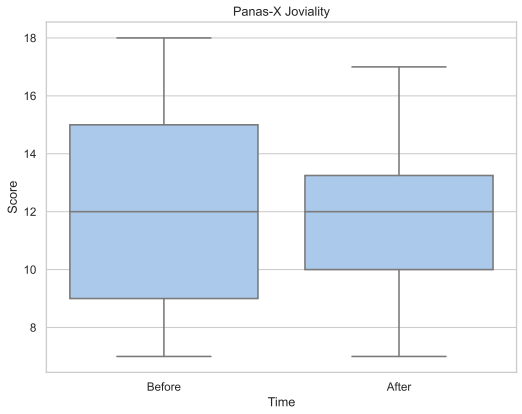
(c) Fear



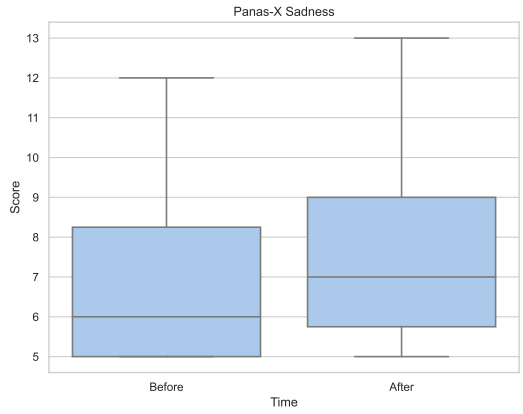
(d) Guilt



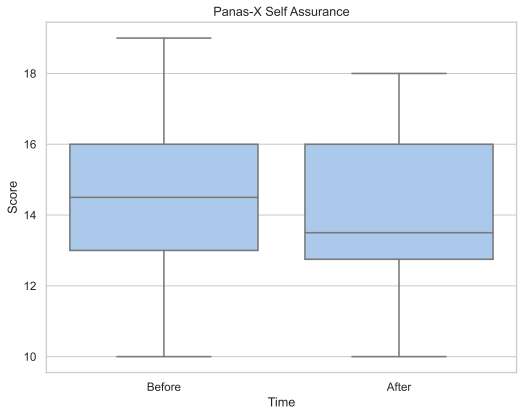
(e) Hostility



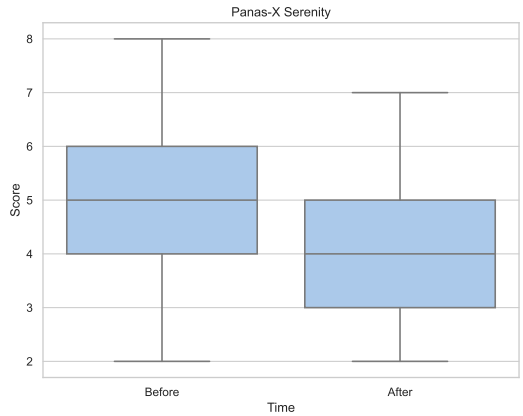
(f) Joviality



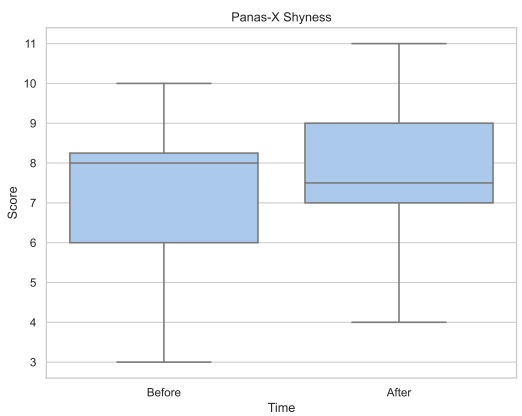
(g) Sadness



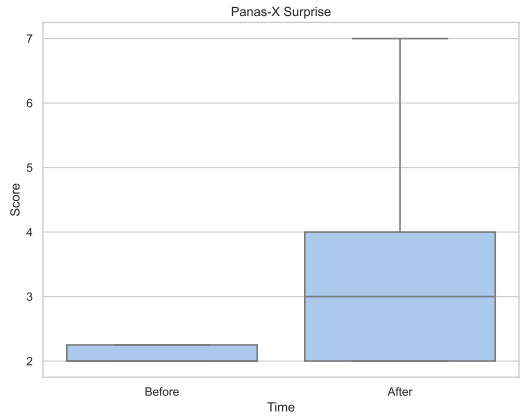
(h) Self Assurance



(i) Serenity

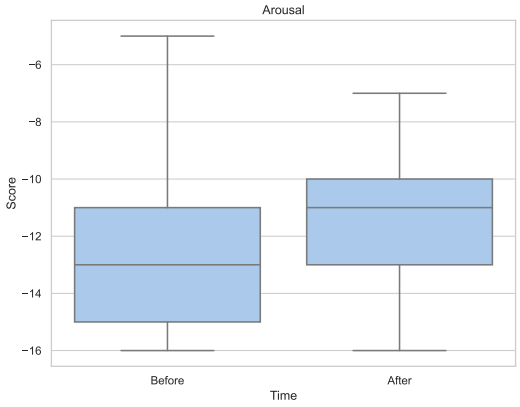


(j) Shyness

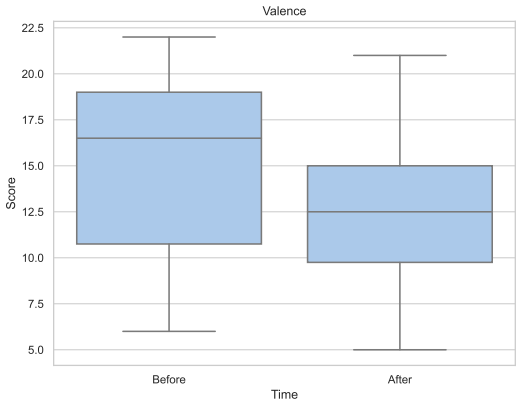


(k) Surprise

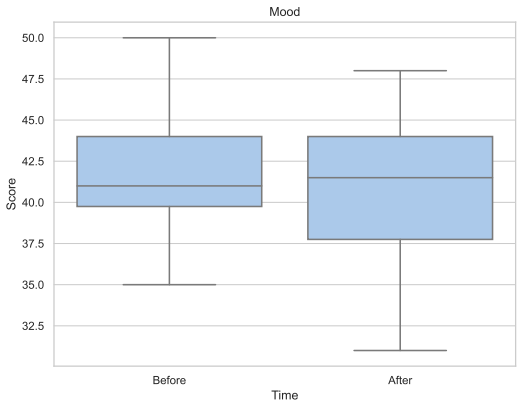
Table C 3: Positive and Negative Affect Schedule-eXpanded (PANAS-X) subscales summary before and after the EEG data collection: (a) Attentiveness, (b) Fatigue, (c) Fear, (d) Guilt, (e) Hostility, (f) Joviality, (g) Sadness, (h) Self Assurance, (i) Serenity, (j) Shyness and (k) Surprise.



(a)



(b)



(c)

Table C 4: Summary of Arousal, Valence, and Mood results before and after the EEG data collection: (a) Arousal, (b) Valence, and (c) Mood.

Appendix D

Code for Section 6.3

D.1 Interpolation map

```
def interpolation_map(value, channels_position, scale=2):
    """
    Interpolate the values of a set of points in a 2D space.

    The function takes as input the values to be interpolated, the coordinates of
        the points
    at which the values are known, and a scaling factor. The scaling factor is used
        to adjust
    the size of the output map.

    The function returns the interpolated map, and the x and y coordinates of the
        interpolated points.

    Parameters
    ----------
    value : array-like
        The values to be interpolated.
    channels_position : dict
        The coordinates of the points at which the values are known.
    scale : int, optional
        The scaling factor. The default is 2.

    Returns
    -------
    interp : numpy.ndarray
        The interpolated map.
    x : numpy.ndarray
        The x-coordinates of the interpolated points.
    y : numpy.ndarray
        The y-coordinates of the interpolated points.
    """

    from scipy import interpolate
    import numpy as np
```

```
# Make a copy of the input arrays and convert them to float32
x = np.array(channels_position['x'], dtype='float32').copy()
y = np.array(channels_position['y'], dtype='float32').copy()

# If all values of x are negative, make them positive
if (np.sum(np.array(x)>=0, axis=0))==0:
    x = np.absolute(x)

# If all values of y are negative, make them positive
if (np.sum(np.array(y)>=0, axis=0))==0:
    y = np.absolute(y)

# Scale the coordinates so that the maximum value is approximately 10
if max(x)<1:
    while max(x)<10:
        x *= 10
        y *= 10
elif max(x)>1:
    while max(x)>10:
        x /= 10
        y /= 10
elif max(x)==1:
    x *= 10**2
    y *= 10**2

# Round the coordinates to the nearest integer and convert them to int32
x = np.around(x/scale, decimals=0)
y = np.around(y/scale, decimals=0)
x = np.array(x, dtype='int32')
y = np.array(y, dtype='int32')

# Shift the coordinates so that the minimum value is 0
x += -min(x)
y += -min(y)

# Get the dimensions of the output map
dimx = max(x)
dimy = max(y)

# Create a list of (x, y) coordinates
pos = list()
for i in range(len(x)):
    pos.append([int(x[i]),int(y[i])])

# Create a grid of points at which to interpolate the values
grid_x, grid_y = np.mgrid[0:dimx,0:dimy]

# Interpolate the values using the cubic method
interp = interpolate.griddata(pos, value, (grid_x, grid_y), method='cubic')

# Rotate the interpolated map by 90 degrees
interp = np.rot90(interp, k=1)
```

```
    return interp , x, y
```

Listing D.1: Compute the topomap with interpolated values

D.2 Data augmentation

```
def data_augmentation(psd_matrix_emotion1, labels_emotion1, num_iterations,
    validation_percentage):
    """
    Performs data augmentation and splits a dataset into training and validation
        subsets using a sliding window with partial overlapping.

    Parameters:
    - psd_matrix_emotion1 (numpy.ndarray): The input PSD matrix for a specific
        emotion.
    - labels_emotion1 (numpy.ndarray): The labels corresponding to the PSD data.
    - num_iterations (int): Number of iterations for data splitting and augmentation
        .
    - validation_percentage (float): Percentage of data to be used for validation.

    Returns:
    - final_psd_matrix (list of numpy.ndarray): Augmented PSD matrices for training.
    - final_labels (list of numpy.ndarray): Corresponding labels for the augmented
        data.
    - validation_psd_matrix (list of numpy.ndarray): PSD matrices for validation.
    - validation_labels (list of numpy.ndarray): Labels for the validation data.
    """
    import numpy as np

    # Calculate the number of samples for validation
    samples_per_iteration = psd_matrix_emotion1.shape[0] / num_iterations
    num_validation_samples = np.ceil(samples_per_iteration * validation_percentage)

    # Generate random indices for validation samples
    random_indices = np.random.randint(0, int(samples_per_iteration), size=int(np.
        floor(num_validation_samples)))

    # Create validation set
    validation_data = []
    for iteration in range(num_iterations):
        for index in random_indices:
            global_index = int(index + iteration * samples_per_iteration)
            validation_data.append(psd_matrix_emotion1[global_index])
            np.delete(psd_matrix_emotion1, global_index)
            np.delete(labels_emotion1, global_index)

    validation_data = np.asarray(validation_data)
    validation_labels = np.full(validation_data.shape[0], 'label_emotion1') #
        Replace 'label_emotion1' with the actual label.

    # Define stopping points for splitting data
```

```
    stopping_points = np.sort(
        np.concatenate([
            np.array([random_indices + x * samples_per_iteration for x in range(
                num_iterations)]).flatten(),
            np.array([samples_per_iteration * x for x in range(num_iterations)]).
                flatten()
        ])
    ).astype('int32')
    stopping_points = np.delete(stopping_points, 0)

    # Split and augment data based on stopping points
    augmented_data = []
    for segment in np.split(psd_matrix_emotion1, stopping_points):
        num_samples = segment.shape[0]
        for i in range(1, num_iterations + 1):
            start_index = segment.shape[2] * segment.shape[3] * i * segment.shape[4]
            end_index = segment.shape[2] * segment.shape[3] * num_iterations * (
                num_samples - 1) * segment.shape[4]
            flattened_array = segment.flatten()[start_index:(end_index + start_index
                )]
            reshaped_segment = np.reshape(
                flattened_array,
                (num_samples - 1, num_iterations, segment.shape[2], segment.shape
                    [3], segment.shape[4])
            )
            augmented_data.append(reshaped_segment)

    augmented_data = np.asarray(augmented_data)
    augmented_data = np.vstack(augmented_data)

    # Assign labels and organize results
    augmented_labels = np.full(augmented_data.shape[0], 'label_emotion1')  # Replace
        'label_emotion1' with the actual label.
    final_psd_matrix = [augmented_data]
    final_labels = [augmented_labels]
    validation_psd_matrix = [validation_data]
    validation_labels = [validation_labels]

    return final_psd_matrix, final_labels, validation_psd_matrix, validation_labels
```

Listing D.2: Data augmentation

D.3 Extraction min set of channels

```
def get_minimum_channels(raw_data, old_montage_files, target_montage_name,
    starting_montage, exclude_channels):
    """
    # This script is designed to work with EEG data that has more than 32 channels.
        Its purpose is to extract a smaller set of 32 channels for data augmentation.
```

```
    1) Focus on the 10-20 standard: The script prioritizes channels that are closest
        to the international 10-20 system for EEG electrode placement.
    This system is widely used and provides a standardized way to compare EEG data
        across different studies.
    2) Iterative selection: The script goes through the available channels step-by-
        step. In each step, it selects the next 32 channels that are closest to the
        10-20 standard.
    3) Continues until exhausted: This process continues until all possible sets of
        32 channels that are close to the 10-20 standard have been extracted.
    4) This script reduce the complexity of your EEG data while ensuring that the
        selected channels are as representative as possible based on the 10-20
        standard.
    This can be useful for various EEG analysis tasks, especially when dealing with
        high-density EEG recordings.

    Identify the minimal set of EEG channels from a starting montage that closely
        approximates a target montage.

    Parameters:
        raw_data: MNE Raw object
            Original raw EEG data.
        old_montage_files: list of str
            List of file paths containing previously computed montages to be
                excluded.
        target_montage_name: str
            The name of the target montage to approximate.
        starting_montage: MNE Montage object
            The starting montage with more channels.
        exclude_channels: list of str
            List of channels to exclude regardless of their position.

    Returns:
        list of str
            Channels to exclude to reduce the starting montage to approximate the
                target montage.
    """

    import numpy as np
    import mne  # MNE library is for EEG data processing

    num_old_montages = len(old_montage_files)

    # Drop non-EEG channels
    channels_to_drop = [ch['ch_name'] for ch in raw_data.info['chs'] if 'EX' in ch['
        ch_name']]

    # Load target montage and extract positions
    target_montage = mne.channels.make_standard_montage(target_montage_name)
    target_positions = target_montage.get_positions()
    target_channel_positions = {
        ch_name: list(pos) for ch_name, pos in zip(target_positions['ch_pos'].keys()
            , target_positions['ch_pos'].values())
    }
```

```
    # Extract starting montage positions
    starting_positions = starting_montage.get_positions()
    starting_channel_positions = {
        ch_name: list(pos) for ch_name, pos in zip(starting_positions['ch_pos'].keys
            (), starting_positions['ch_pos'].values())
    }

    # Determine channels present in the starting montage
    available_channels = list(starting_channel_positions.keys())

    # Find the minimal channels to approximate the target montage
    channel_distances = {}
    selected_channels = []
    for target_ch, target_pos in target_channel_positions.items():
        min_distances = {}
        for start_ch, start_pos in starting_channel_positions.items():
            if (start_ch in available_channels and
                start_ch not in old_montage_files and
                start_ch not in channel_distances.values()):
                distance = np.sqrt(np.sum(np.square(np.asarray(target_pos) - np.
                    asarray(start_pos))))
                min_distances[f"{target_ch}_{start_ch}"] = distance
        if min_distances:
            min_val = min(min_distances.values())
            selected_channel = next(ch for ch, dist in min_distances.items() if dist
                == min_val)
            selected_channels.append(selected_channel)

    # Build a dictionary mapping target to selected channels
    channel_distances = {
        item.split('_')[0]: item.split('_')[1] for item in selected_channels
    }

    # Prepare the list of channels to exclude
    new_distances = {ch: starting_channel_positions[ch] for ch in channel_distances.
        values()}
    excluding_channels = list(set(starting_channel_positions.keys()) - set(
        new_distances.keys()))
    excluding_channels = list(set(excluding_channels) - set(available_channels)) +
        channels_to_drop

    return excluding_channels
```

Listing D.3: Channel set extraction

Appendix E

Equations

N.	Description	Formula
(3.1)	Average noise in ERP	$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} x_i^c(t) = 0$
(3.2)	Artifact regression	$EEGcor = EEGraw - \gamma Factor(ArtifactChannel)$
(3.3)	DWT	$\psi_{j,k}(t) = \frac{1}{\sqrt{a_0^j}} \psi\left(\frac{t - k b_0 a_0}{a_0^j}\right)$
(3.4)	BSS	$X = AS$
(3.5)		$U = WX$
(3.6)	EMD	$\alpha(t) = \frac{\epsilon_{max}(t) + \epsilon_{min}(t)}{2}$
(3.7)		$d(t) = s(t) - \alpha(t)$
(3.8)		$g(t) = s(t) - IMF_1(t)$
(3.9)		$s(t) = \sum_{i=1}^{M} IMF_i(t) + R(t)$
(3.10)	FT	$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn}, \quad k = 0, 1, \ldots, N-1$
(3.11)	PSD	$\mathrm{PSD}(f) = \frac{1}{K} \sum_{i=1}^{K} \frac{1}{L} \left|\mathrm{FT}_i(f)\right|^2$
(3.12)	Output of neuron in ANN	$y = f(w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + b)$
(3.13)	Nonlinear activation function	$f(z) = \frac{1}{1+e^{-z}}$
(3.14)	Compute error	$\frac{\partial J}{\partial w_i} = \frac{\partial J}{\partial y} \cdot \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial w_i}$
(3.15)	Update neuron	$w_i \leftarrow w_i - \eta \frac{\partial J}{\partial w_i}$
(3.16)	Fully connected layer	$y = f(w_1 x_1 + w_2 x_2 + \cdots + w_n x_n + b)$
(3.17)	Convolutional layer	$Y[i,j] = \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} X[i-k_1, j-k_2] W[k_1, k_2]$
(3.18)	Hidden state in RNN	$h_t = f(W_h h_{t-1} + W_x x_t + b)$
(3.19)	Output at each instant RNN	$y_t = g(W_y h_t + c)$
(3.20)	Updating LSTM	$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$
(3.21)		$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$
(3.22)		$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$
(3.23)		$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$
(3.24)		$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$
(3.25)		$h_t = o_t \odot \tanh(C_t)$
(3.26)	Accuracy	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
(3.27)	Precision	$Precision = \frac{TP}{TP+FP}$
(3.28)	Recall	$Recall = \frac{TP}{TP+FN}$
(3.29)	F1 score	$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$
(3.30)	Ideal filter	$H(f) = \begin{cases} 1, & f_1 \le f \le f_2, \\ 0, & f < f_1 \text{ o } f > f_2. \end{cases}$
(5.1)	Band power	$BandPower_i = \sum_{b_i f=b_i f_{min}}^{b_i f_{max}} P_{xx}(f)$
(5.2)	Total power	$TotalPower = \sum_{f=f_{min}}^{f_{max}} P_{xx}(f)$
(5.3)	Relative band power	$RelativeBandPower = \frac{BandPower_i}{TotalPower}$

Continued on next page

Table E 1 – continued from previous page

N.	Description	Formula
(6.1)	Graphs	$G = (V, E)$
(6.2)	Complex signal	$z(t) = x(t) + i\widetilde{x}(t)$
(6.3)	HT	$\widetilde{x}(t) = \frac{1}{\pi} PV \int_{-x}^{\infty} \frac{x(\tau)}{t-\tau} d\tau$
(6.4)	IP	$\phi(t) = arctan \frac{\widetilde{x}(t)}{x(t)}$
(6.5)	IP difference	$\phi_{i,j}(t) = \phi_i(t) - \phi_j(t)$
(6.6)	PLI	$pli_{i,j} = \lvert \frac{1}{T} \sum_{t=1}^{T} \text{sign}(\phi_{i,j}(t)) \rvert$
(6.7)	Sign operator	$\text{sign}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$
(6.8)	Autocorrelation	$z[k] = \frac{\sum_{l=0}^{\lVert x \rVert - 1} x_l y^*_{l-k+N-1}}{\sqrt{(x \cdot x)(y \cdot y)}}$
(6.13)	Correlation maps	$time \times height_map \times length_map \times band$

Table E 1: Equations.

Part VI

Lists

List of Figures

1.1 This manuscript addresses topics of increasing complexity, both in terms of scientific consensus and the absence of a ground truth. This increasing complexity culminates in the domain of emotions, where it is impossible to have a definite ground truth in emotional datasets. This is due to the high inter-subject variability and the instability of emotions within the same subject, as well as the lack of a universally accepted theory of emotions within the scientific community. 2

2.1 A Cajal's paint representing neurons. Image from the "*Cajal Institute*" and the "*Spanish National Research* Council". 6
2.2 An astrocyte. (Author: "GerryShaw", CC-SA). 7
2.4 An overview of the CNS. Image from [14]. 8
2.3 Multipolar neuron with its several components. (Author: "BruceBlaus", CC-BY). . . 8
2.5 The anatomical planes. (Author: "Zwarck", CC-SA). 9
2.6 The Brodmann areas. Image from [15]. 10
2.7 The brain structures. Creative Commons license. 11
2.8 Cortex's layers and EEG current generation. Image from [16]. 12
2.9 Action potential diagram. Image from [17]. 12
2.10 Currents detected by EEG. Image from [18]. 12
2.11 Thalamo-cortico-thalamic circuits. Image from [19]. 13
2.12 Motor pathways. Image from [21]. 14
2.13 The somatomotor and somatosensory homunculus. Image from [22]. 15
2.14 The sleep architecture with some elements. Image from [26]. 16
2.15 Elements of the sleep stage NREM 2. (Author: "Neocadre", Public Domain.) 16
2.16 The ARAS network. Image from [29]. 17
2.17 The two model sleep stage. Image from [35]. 18
2.18 The three model sleep stage. Image from [36]. 18
2.19 The Circumplex Model with Arousal and Valence dimensions 20
2.20 3D Russel's emotional space with Arousal, Valence, and Dominance dimensions (the 2D version of the space contains only Arousal and Valence). 20
2.21 The neural circuits involved in emotional process. Image form [42]. As can be seen, the amygdala (AMY) is a small structure that is closely connected to the hippocampus and parahippocampus (HPC/PHPC), which is responsible for memory consolidation processes. This partly explains how memories with a high emotional content are remembered better than others. 21

2.22 Emotional pathways. Image from [43]. 22

3.1 The three different ways to record electrical activity from the brain. Image from [62]. 24
3.2 The MEG instrument. (Author: "NIMH author", Public Domain). 25
3.3 A brief description of how MEG works. Image from [64]. 25
3.4 (left) A structural MRI and (right) an fMRI acquistion on the MRI after the skull removing. Image from [69] 27
3.5 An EEG-fNIRS instrument. Image from BrainProducts®website. 27
3.6 EEG data. On y-axis, for each row (channels) the amplitude in μV, on x-axis, the time in seconds. The spikes are eye movement artifacts. 28
3.7 Standardized position of electrodes. Image from [74]. 29
3.8 P300 wave. Image from [77]. 30
3.9 Example of Topo-ERP from Chapter 6. 31
3.10 Topographical maps constructed on the basis of ICs calculated by ICA. Image from [83]. BEOG: blink (a); VEOG: vertical eye movement (b); HEOG: horizontal eye movement (c); ECG: cardiac artifact (d); EMG: muscular artifact (e); IF: high impedance (f). 32
3.11 Example of ICA decomposition results: (a) raw EEG signal, (b) ICA sources, (c) example of one IC with its properties. The different background colours in (a) and (b) represents different mental tasks (green: rest; blue: feet movement) 34
3.12 An example of the EMD decomposition: (a) original EEG signal, (b) local maxima and minima, (c) average computing, (d) the process is repeated until the function obtained is a monotonic function, and the results are the IMFs. 35
3.13 Time-frequency representations using **(a)** FT and **(b)** MvFIF showed though IMFogram of EEG data from DEAP dataset [97]. 36
3.14 Example of CSP decomposition results: (a) raw EEG signal, (b) CSP components and (c) CSP patterns. The different background colors represent different mental tasks (green: rest; blue: feet movement) 36
3.15 The power spectrum of EEG signal after applying a bandpass FIR filter between 7 and 30 Hz. On the upper right corner, the electrodeds position coded with the same color of each spectrum of the image. The dotted line represent the bandpass filter. . . 38
3.16 Example of time-frequency representation(ERSP) of Fp1 channel in two different experimental condition: on the left the condition H, in the center the condition S, and on the right the result of the p-value for t-test between the two conditions. The data cames from the experiment shown in Chapter 6. 39
3.17 Example of topomap representation called Topo-ERSP. The data cames from the experiment shown in Chapter 6. 39
3.18 A graph-based visualization of the EEG FC based on the PLI. The data cames from the experiment shown in Chapter 6. Image generated with EEGraph library [112]. . . 40
3.19 The sigmoid function. 42
3.20 A diagram of a classic ANN. Image from [119]. 43
3.21 ANN diagram. (Author: "Chrislb", CC-SA). 44
3.22 CNN diagram. (Author: "Aphex34", CC-SA). 44

3.23 Diagram of an LSTM cell. 46
3.24 The Conv-LSTM layer structure. Image from [123]. 46
3.25 EEGnet Architecture. Image from [124]. 48
3.26 EEG-ITnet Architecture. Image from [127]. 49
3.27 Deep4Net Architecture. Image from [128]. 50
3.28 An ideal BCI that contains both *passive* and *active* mental-state detection. 51
3.29 A standard pipeline for EEG data analysis. 51
3.30 Example of (a) raw EEG signal and (b) filtered signal between 1 a 30 Hz. The different background colours in (a) and (b) represents different mental tasks (green: rest; blue: feet movement) 52
3.31 A band-pass filter. Image from [139]. 53
3.32 (a) raw EEG signal and (b) the EEG signal after PREP cleaning process. 54

4.1 A picture of the VG. 57
4.2 The pipeline used for data preparation and classification stages. 60
4.3 The split of the original set in train, validation and test set after preprocessing and epochs extraction. 61
4.4 *Network Tree* - the first classification of the signal is performed by the $EEGnet_{LR}$ and then a classification check is carried out. If L is identified, the input signal is sent to the $EEGnet_{0L}$ for final classification between 0 or L. In case R is classified, the input signal is sent both to the $EEGnet_{0L}$ and $EEGnet_{0R}$. The result of their combined classification returns the result of the final classification between 0 or R. The input undergoes, before being sent to a network, a transformation by the associated CSP. . 62
4.5 In the upper part, the pipeline applied to clean the signals is shown performed for each participant. The training and test phases are shown in the lower part after merging all samples in one set. The processes in the blue box are skipped if the CSP transformation is not applied. 65
4.6 The channels used in this section. The channels are located in the retral part of the frontal area, on the central area and the anterior part of the parietal area, covering the motor areas. This process was done due to reduce the useless information not related to sensory-motor areas, improving the learning process. 66
4.7 The confusion matrix of the EEGnet without CSP transform for both datasets, **(a)** the PhysioNet dataset and **(b)** the Upper Limb dataset. 69
4.8 EEG raw data before (**a**) and after (**b**) a cleaning band-pass [1–45 Hz] filtering. In (**a**), the high noise level makes it difficult to visualize high-amplitude brain signals. Vertical lines represent the triggers for rest, fixation cross, and open/close movement, respectively. 78
4.9 Model based on data collected by VG. The left part shows the vertical view; the right part shows the horizontal view. The Cartesian reference system is reported in the centre. 79
4.10 Acquisition environment scheme (**a**) and real acquisition environment (**b**) used in this section. 79

4.11 The analyzed movements. The dotted line represents the movement acquired but not analyzed in this section. The class "movement" is created by merging the open/close and wrist rotation classes. . . . 79
4.12 The used hand movement protocol. . . . 80
4.13 Instructions shown to participants for both MI and ME. The flow stops when 8 repetitions are reached. . . . 80
4.14 Preprocessing for sample generation and training. The colored boxes represent the parameters that change to explore different frequency bands (in orange) and the baseline reduction method (in pink), for each combination of movement/rest. The movement class (violet box) represents the merging of the open/close and wrist rotation. In the lower part of the scheme, the training process of the model is described. . . . 81
4.15 x (**a**), y (**b**), and z (**c**) components of the fingertip trajectory for the Horizontal LMC. All fingertips are reported in a single plot. . . . 82
4.16 x (**a**), y (**b**), and z (**c**) components of the fingertip trajectory for the Vertical LMC. All fingertips are reported in a single plot. . . . 83
4.17 x (**a**), y (**b**), and z (**c**) components of the fingertip velocity for the Horizontal LMC. All fingertips are reported in a single plot. . . . 84
4.18 x (**a**), y (**b**), and z (**c**) components of the fingertip velocity for the Vertical LMC. All fingertips are reported in a single plot. . . . 85
4.19 Spatial distribution of PSD for each task in a single subject. The timeline is the same as used in Figures 4.15–4.18. . . . 86

5.1 Use case of ASIS. . . . 89
5.2 The 10-20 system electrode positions. The four channels with the same color record two signals based on two pairs of electrodes in a bipolar configuration. . . . 89
5.3 Feature extraction flow-chart used in this chapter. . . . 92
5.4 Dataset generator. . . . 93
5.5 System classification and prediction. In this scheme, an artifact rejection algorithm is included because the ASIS is conceptualized for a real-time application [83, 85]. . . . 94
5.6 Description of the ASIS system. The prediction timeline stats its label forecasting and upon reaching the conclusion of the Alarm Clock Zone, the system rouses the subject at the last phase of light sleep to extend sleep duration and ensure an optimal wake-up. This suggests that the system delays waking until the most advantageous stage within the Alarm Clock Zone is achieved, if multiple appropriate stages exist. When the designated wake-up time arrives, the classifier assesses whether the current stage is appropriate for awakening. Should no appropriate stages be available, the alarm is set off at the maximum boundary of the Alarm Clock Zone. Note: the array sizes depicted do not match 30-second intervals on the timeline; they are merely illustrative of the procedure. . . . 95
5.7 LSTM network architecture . . . 97
5.8 Confusion matrix for prediction stage. . . . 97

6.1 Trend of the number of publications by searching, in the PubMed database, the terms *"Emotion" AND "Recognition"* in *"article title" OR "abstract"* (searching date: January 1st, 2025). . . . 101
6.2 EEG brain sensitivity map (in yellow), with sensitivity decreasing with depth, and the main circuits of the limbic system, where emotions mostly originate (in red). . . . 104
6.3 Experimental Protocol for the IMAGINED EMOTIONS dataset. . . . 107
6.4 Experimental protocol for DEAP dataset. . . . 109
6.5 Experimental Protocol for the "Emotion recognition experiment" of MAHNOB dataset. 110
6.6 Experimental Protocol for the "Implicit taggging experiment" of MAHNOB-HCI dataset. 110
6.7 Experimental Protocol for the SEED dataset. . . . 110
6.8 Experimental Protocol for the ASCERTAIN dataset. . . . 111
6.9 Experimental Protocol for the DREAMER dataset. . . . 112
6.10 Experimental Protocol for the SEED-IV dataset. . . . 112
6.11 Experimental Protocol for the AMIGOS dataset. . . . 114
6.12 Experimental Protocol for the SEED-V dataset. . . . 114
6.13 Experimental Protocol for the MUSEC dataset. . . . 115
6.14 A graph of electrodes connection for β band of "happiness" emotion by subject n. 10. The strength of PLI for each pair of nodes is shown by the thickness and the color intensity of the line: higher thickness and more red represent an higher value of PLI. It can be noticed not all channels were included due to original authors exclusion of bad channels in the dataset. . . . 118
6.15 Simulation of PLI between four electrodes. . . . 119
6.16 Positions of the residual common electrode (133 channels). . . . 120
6.17 3D tensor: correlation maps of one epoch generated for the six frequency bands. . . . 125
6.18 4D tensor: temporal sequence of 3D tensors. The resulting 4D tensor is used as the input of the LSTM. . . . 125
6.19 4D tensor production flow-chart. . . . 126
6.20 General scheme of the used architecture. In particular, the 4D tensor input and the binary classification output are reported. The internal design of the LSTM network is reported in Fig. 6.21 . . . 126
6.21 LSTM architecture in details. . . . 127
6.22 Model *accuracy* during training (blue line) and validation (orange line). . . . 127
6.23 Model *loss* during training (blue line) and validation (orange line). . . . 128
6.24 Dimensional space of the NAPS labels. . . . 133
6.25 The image selection process . . . 134
6.26 (a) The paradigm and (b) environment of the first part of the experiment. The questionnaires are explained in Appendix B. . . . 137
6.27 (a) The paradigm and (b) environment of the second part of the experiment. The questionnaires are explained in Appendix B. . . . 138
6.28 The preprocessing. . . . 139
6.29 Results for statistical test with $p-value < 0.05$ of ERP. The time interval is between -200 ms and +600 ms for HNS condition. (a) Frontal left, (b) Frontal right, (c) Central left, (d) Central right, (e) Parietal left, (f) Parietal right. . . . 141

6.30 Results for statistical test with $p-value < 0.05$ of ERSP. The time interval is between -200 ms and +600 ms for HNS condition, and the frequencies analysed goes from 1 to 100 Hz. (a) F left, (b) F right, (c) C left, (d) C right, (e) P left, (f) P right. 142
6.31 Topographical ERSP representation and statistical test between "happiness", "neutral" "sadness" (HNS) conditions in the following bands (a) theta, (b) alpha, (c) beta and (d) low gamma. The red electrodes indicates $p - value < 0.05$. The time interval is between 0 an 150 ms. 144
6.32 Topographical ERSP representation and statistical test between "happiness", "neutral" "sadness" (HNS) conditions in the following bands (a) theta, (b) alpha, (c) beta and (d) low gamma. The red electrodes indicates $p - value < 0.05$. The time interval is between 150 an 300 ms. 145
6.33 Topographical ERSP representation and statistical test between "happiness", "neutral" "sadness" (HNS) conditions in the following bands (a) theta, (b) alpha, (c) beta and (d) low gamma. The red electrodes indicates $p - value < 0.05$. The time interval is between 300 an 450 ms. 146
6.34 Confusion matrices results for Comparison 2 (Image VS Blank screen) for (a) KNN, (b) linear SVM and (c) K-means. 147
6.35 The (a) confusion matrix and (b) accuracy and (c) loss on validation set during the training with Image vs Blank Screen classes. 147
6.36 An example of the interaction between memory process (*baseline*) and the emotions on frequency-channel plane. Each emotion has a specific areas in which, with more or less intensity coded by color, it act at specific place and frequency. The crossing between the baseline and the emotions represent the most difficult point of recognition due to the interference. The colours in the graph have been made more transparent than in the color bar to highlight any overlaps between the different ribbons. 150
6.37 An example of the spatial distribution of the emotion used in REMEMO on Russell Circumplex Plane. The transparent dots represent the theoretical place of the emotions, while the translated dots represent the emotions recorded that are translated by ϵ depending on the mood at the time of recording. The fixation cross, with no emotional relevance, is translated only by the value of the mood. The red arrow represents the translation vector, which is the mood of the participant at the recording time. 151
6.38 Experimental Protocol for the REMEMO dataset. The questionnaires are explained in Appendix B. 151
6.39 Environment isolated from noise and external stimuli with a uniformly lighting. . . . 152
6.40 The pipeline used for preprocessing. 152
6.41 (a) ERP for centro-parietal, between -200 ms and 800 ms, (b) Topo-ERP between 0 ms and 800 ms. 153
6.42 Topo-ERSP for whole period excluding the first second avoiding the influence of the reading, for frequency band between 4 - 8 Hz. The threshold for significance is $p - value < 0.05$. 153

6.43 Topo-ERSP for whole period excluding the first second avoiding the influence of the reading, for frequency band between 31 - 48 Hz. The threshold for significance is $p - value < 0.05$. 154

List of Tables

2.1 Structions and functions of cerebral cortex [20]. 14
2.2 EEG Frequency Bands and Sleep-related Events [24]. 16

3.1 Binary confusion matrix with TP, TN, FP), and FN 47
3.2 Multi-class confusion matrix with TP, FP), and FN for each class. 47

4.1 The hyperparameters used to train all EEGnet 61
4.2 Prediction metrics table for all networks used in *Network Tree* 62
4.3 Results of *Network Tree* 63
4.4 Parameters for each IC-label category. The *"N"* letters mean that these components are not removed, while the other values represent the threshold to choose between leaving and removing the component if it was classified as the label. 65
4.5 The hyperparameters used to train all Network 66
4.6 The validation accuracy and the training and validation loss curves are plotted for the EEGnet. The best results were achieved from the EEGnet trained without CSP transformation, for both datasets. The blue line indicates the change in batch size during the learning process from 16 to 32, and the green line indicates when the early stopping function occurs. The first two rows illustrate the curve related to the PhysioNet dataset, whereas the subsequent two rows show the curve for the Upper Limb dataset, without and with CSP. Each column represents Training Loss, Validation Loss, and Validation Accuracy respectively for each training session. 67
4.7 Prediction metrics table for all networks used in *PhysioNet* and *Upper Limb* dataset without CSP. In bold, the best accuracy for each dataset and model. 68
4.8 Prediction metrics table for all networks used in *PhysioNet* and *Upper Limb* dataset with CSP. In bold, the best accuracy for each dataset and model. 68
4.9 The EEGnet hyperparameters. 74
4.10 Overview of the results. The movement class represents the merging of open/close and wrist rotation epochs. The best results for each combination of *baseline* and *filter* for each binary classification are in the box. In orange are the results that go above the chance threshold [145]. The box around the value indicates the best results for the classification. 75
4.11 Results performing 5-fold cross-validation on the best models reported in Table 4.10. 76

5.1 Sleep Stages and Descriptions 88
5.2 Comparison between different classification systems [28]. 90

5.3 Probability of occurrence of each class in AASM classification. 94
5.4 Sample number of the training set for LSTM. 96
5.5 Performance on Prediction with LSTM on AASM labels. 96
5.6 Metrics for Prediction with LSTM. 97
5.7 Distribution of the stages in the dataset. 97
5.8 List of features used. 98

6.1 Factor underline emotion related activation [273]. 100
6.2 Comparison between all datasets 106
6.3 Emotions analyzed in IMAGINED EMOTIONS dataset. 107
6.4 Emotional Evaluation for the "Emotion recognition experiment" in MAHNOB-HCI. . 109
6.5 Emotions analyzed in MAHNOB-HCI database. 109
6.6 Emotions analysed in DREAMER database. 111
6.7 Emotions in SEED-IV dataset 112
6.8 Emotions analyzed in AMIGOS dataset. 113
6.9 Connectomes of four emotions located as in the Russel plane. For each emotion, the connectomes are divided by subjects and by frequency bands. The PLI value is coded by the thickness and the color of the edge. Under the table, color palette is reported. The PLI value goes from 0 (yellow) to 1 (blue). 122
6.10 Connectome averaged by subjects, separated for emotions and frequency bands. The value PLI is encoded by the thickness and the color of the edge. In the table, the color palette is reported. The value of PLI goes from 0 (yellow) to 1 (blue). 122
6.11 Sum of PLI for each subject filtered only with 1-49 Hz band-pass filter, separated for emotions and subjects 122
6.12 Sum of PLI, separated by subject, by emotions and by frequency bands (δ, θ, α, β, γ are represented in the horizontal axis of each graph). Sum of PLI values are the vertical axis of each graph. 123
6.13 Accuracy values, separated by subject, obtained during the test when recognizing "relax" and "disgust". Data from the 4 subjects shown in the gray background were not used for training and validation. 128
6.14 Three dimension model used in NAPS-BE labeling process. 132
6.15 Study designs. In *italic* the design used in this section. 140
6.16 Analysis performed on the data. 140
6.17 Results for Comparison 2 143
6.18 Metrics for Image vs Blank Screen classes experiment with EEGNet 148
6.19 Emotions categorized by Valence and Arousal. 150

A 1 Shapiro-Wilk Statistic normality test on RBP distribution for each band power and each electrode. 188
A 2 Histograms of the RBP average of sleep stage divided per channels and frequency bands. 189
A 3 Confusion matrix of the Mann-Whitney test performed on the distribution of the RBP of sleep stage divided per channels and frequency bands. 190
A 4 Histograms of the RBP average of sleep stage divided per stage and frequency bands. 191

A 5 Confusion matrix of the Mann-Whitney test performed on the distribution of the RBP of sleep stage divided per stage and frequency bands. 192
A 6 Distribution of RBP of sleep stage divided per channels and frequency bands. 193
A 7 Distribution of RBP of sleep stage divided per channels and sleep stage. 194
A 8 Normalized distribution of RBP of sleep stage divided per channels and frequency bands. 195
A 9 Normalized distribution of RBP of sleep stage divided per channels and sleep stage. . 196
A 10 Power spectrum for each channel and sleep stage in AASM classification. 197

C 1 Brief Mood Introspection Scale (BMIS) results summary before and after the EEG data collection: (a) BMIS Arousal Calm, (b) BMIS Negative Relaxed, (c) BMIS Pleasant Unpleasant and (d) BMIS Positive Tired. 200
C 2 PANAS-X results summary before and after the EEG data collection: (a) Negative Affect and (b) Positive Affect. 201
C 3 Positive and Negative Affect Schedule-eXpanded (PANAS-X) subscales summary before and after the EEG data collection: (a) Attentiveness, (b) Fatigue, (c) Fear, (d) Guilt, (e) Hostility, (f) Joviality, (g) Sadness, (h) Self Assurance, (i) Serenity, (j) Shyness and (k) Surprise. 203
C 4 Summary of Arousal, Valence, and Mood results before and after the EEG data collection: (a) Arousal, (b) Valence, and (c) Mood. 204

E 1 Equations. 212

List of Algorithms

1 Feature Extraction Algorithm 92
2 LSTM Dataset Generator Algorithm 92
3 System overview 93
4 Connectome Generator 119
5 Image selection Algorithm from NAPS-BE and SFIP 135

Acronyms

AASM American Academy of Sleep Medicine

AC Affective Computing

AI Artificial Intelligence

ALIF Adaptive Local Iterative Filtering

ANN Artificial Neural Network

ANOVA Analysis of Variance

ANS Autonomic Nervous System

ARAS Ascending Reticular Activating System

ASR Artifact Subspace Reconstruction

BCI Brain-Computer Interface

BDI Beck Depression Inventory

BEM Boundary Element Method

BMIS Brief Mood Introspection Scale

BOLD Blood-Oxygen-Level-Dependent

BSS Blind Source Separation

CAR Common Average Referencing

CCA Canonical Correlation Analysis

CNN Convolutional Neural Network

CNS Central Nervous System

CNV Contingent Negative Variation

ConvLSTM Convolutional LSTM Neural Network

CPV Cauchy Principal Value

CSP Common Spatial Pattern

CV Computer Vision

DEAP Dataset for Emotion Analysis using Physiological signals

Deep4Net Deep 4-Net

DICOM Digital Imaging and COmmunications in Medicine

DL Deep Learning

DWT Discrete Wavelet Transform

EC Effective Connectivity

ECG Electrocardiography

ECoG Electrocorticography

EEG Electroencephalography

EEGITNet EEG Inception Temporal Network

EEGNet EEGNet

EEMD Ensemble Empirical Mode Decomposition

eLORETA Exact Low Resolution Electromagnetic Tomography

EMD Empirical Mode Decomposition

EMG Electromyography

EOG Electrooculography

EP Evoked Potential

ERD Event-Related Desynchronization

ERM Entropy Rate Minimization

ERP Event-Related Potential

ERP N1 N1 Event-Related Potential

ERP P2 P2 Event-Related Potential

ERS Event-Related Synchronization

ERSP Event-Related Spectral Perturbation

ESP Event-Stimulus Potential

ETR Electroencephalography Topographical Representation

FC Functional Connectivity

FEM Finite Element Method

FID Free Induction Decay

FIF Fast Iterative Filtering

FIR Finite Impulse Response

fMRI functional Magnetic Resonance Imaging

FN False Negative

fNIRS functional Near-infrared Spectroscopy

FP False Positive

FT Fourier Transform

GC Granger Causality

GPDC Generalized Partial Directed Coherence

HCI Human-Computer Interaction

HRI Human-Robot Interaction

HT Hilbert Transform

IAPS International Affective Picture System

IC Independent Component

ICA Independent Component Analysis

iEEG Intracranial Electroencephalography

IMF Intrinsic Mode Function

IOD Information Object Definition

IP Instantaneous Phase

IR Infrared

ITI Inter-Trial Interval

K-means K-means

KNN K-nearest Neighbors

LDA Linear Discriminant Analysis

LMC Leap Motion Controller

LORETA Low Resolution Electromagnetic Tomography

LOSOCV Leave-One-Subject-Out Cross-Validation

LPP Late Positive Potential

LR Learning Rate

LSL Lab Streaming Layer

LSTM Long Short Term Memory

M1 Primary Motor Cortex

MCMI-III Millon Clinical Multiaxial Inventory

MDD Major Depressive Disorder

ME Motor Execution

MEG Magnetoencephalography

MEMD Multivariate Empirical Mode Decomposition

MI Motor Imagery

ML Machine Learning

MLP MultiLayer Perceptron

MMN Mismatch Negativity

MRCP Motor-Related Cortical Potential

MRI Magnetic Resonance Imaging

MS Multiple Sclerosis

MSE Mean Squared Error

MvFIF Multivariate Fast Iterative Filtering

N170 N170 Event-Related Potential

N400 N400 Event-Related Potential

NAPS Nencki Affective Picture System

NAPS-BE Nencki Affective Picture System Basic Emotion

NB Naive Bayes

NBC Node Betweenness Centrality

NREM Non-Rapid Eye Movement

OHBDS Open Hemispheric Brain Dominance Scale

P200 P200 Event-Related Potential

P300 P300 Event-Related Potential

PANAS Positive and Negative Affect Schedule

PANAS-X Positive and Negative Affect Schedule-eXpanded

PC Principal Component

PCA Principal Component Analysis

PCC Pearson Correlation Coefficient

PGO Ponto-geniculo-occipital waves

PhB Phase-based (EEG)

PHQ-9 Patient Health Questionnaire-9

PLI Phase Lag Index

PLV Phase Locking Value

PMA Premotor Area

PNS Peripheral Nervous System

PoB Power-based (EEG)

PREP early stage PREProcessing

PSD Power Spectral Density

PSQI Pittsburgh Sleep Quality Index

QPR Quantitative Physical Rehabilitation

RBF Radial Basis Function

RBP Relative Band Power

REM Rapid Eye Movement

REMEMO REMembrance of EMOtions

REST Reference Electrode Standardization Technique

RF Radiofrequency

RK Rechtschaffen and Kales

RNN Recurrent Neural Network

SAM Self-Assessment Manikin

SEED SJTU Emotion EEG Dataset

SFIP Set of Fear Inducing Pictures

SGD Stochastic Gradient Descent

sLORETA standardized Low Resolution Electromagnetic Tomography

SMA Supplementary Motor Area

SNR Signal-to-Noise Ratio

SNS Sympathetic Nervous System

SSVEP Steady-State Visually Evoked Potential

STAI State-Trait Anxiety Inventory

STAI-S STAI-State

STAI-T STAI-Trait

SVM Support Vector Machine

SWA Slow-Wave Activity

SWI Small-World Index

SWS Slow-Wave Sleep

t-test t-test

T1 T1 relaxation time

T2 T2 relaxation time

TAS-20 Toronto Alexithymia Scale

TE Transfer Entropy

TN True Negative

Topo-ERP Topographical Event-Related Potential

Topo-ERSP Topographical Event-Related Spectral Perturbation

TP True Positive

VG Virtual Glove

VR Virtual Reality

WT Wavelet Transform

Nomenclature

Accuracy The proportion of correctly classified instances out of the total number of instances.

Actigraph A device equipped with a gyroscope that measures the subject's movements.

Adaptive Local Iterative Filtering A signal decomposition technique that separates a non-stationary signal into its constituent components, called Intrinsic Mode Functions (IMFs), by iteratively filtering the signal with adaptive and data-driven filter lengths. It is particularly effective for analyzing signals with amplitude and frequency modulation.

Affective Computing A field of computer science and engineering that aims to develop systems and devices that can recognize, interpret, process, and simulate human emotions. It involves the design and development of algorithms, sensors, and systems that can detect, analyze, and respond to emotional cues, such as facial expressions, voice intonation, and physiological signals.

Alpha waves Brainwaves in the 8-12 Hz range, present during relaxed wakefulness and light meditation.

American Academy of Sleep Medicine A professional organization dedicated to the advancement of sleep medicine and research. The AASM has updated and refined the Rechtschaffen and Kales sleep scoring criteria, providing the current gold standard for sleep stage classification. The sleep stages identified were three NREM stages (N1, N2, N3) and one REM stage.

AMIGOS A dataset for Multimodal research of affect, personality traits and mood on Individuals and GrOupS. A multimodal dataset for affect, personality, and mood research. It includes EEG, ECG, GSR, and video data from participants watching emotional videos individually and in groups. AMIGOS is unique for its inclusion of both short and long videos and its focus on individual and group dynamics.

Analysis of Variance A statistical test used to compare the means of two or more groups. It assesses whether there are statistically significant differences between the group means by analyzing the variance within and between the groups.

Anti Sleep Inertia System System designed to determine the best time to wake up using an alarm clock within a specified time window to reduce the sleep inertia, improve the sleep time and optimize the sleep session.

Arousal The state of being physiologically alert, awake, and attentive.

Artifact Subspace Reconstruction A method for removing artifacts from EEG data by identifying and reconstructing the subspace of the data that is dominated by artifacts. It uses a sliding window approach to identify segments of data that are likely to contain artifacts and then projects these segments onto a clean data subspace. ASR is effective at removing a wide range of artifacts, including eye blinks, muscle activity, and line noise.

Artificial Intelligence The ability of a computer or a robot controlled by a computer to perform tasks that are usually done by humans because they require human intelligence and discernment.

Artificial Neural Network A computational model inspired by the structure and function of biological neural networks. It consists of interconnected nodes (neurons) organized in layers, which process and transmit information through weighted connections.

Ascending Reticular Activating System A set of connected nuclei in the brains of vertebrates that is responsible for regulating wakefulness and sleep-wake transitions. Its main function is to modify and potentiate thalamic and cortical function such that electroencephalogram (EEG) desynchronization ensues.

Autonomic Nervous System The part of the nervous system that controls involuntary bodily functions, such as breathing, heart rate, digestion, and sweating. It is divided into two branches: the sympathetic nervous system, which prepares the body for "fight or flight" responses, and the parasympathetic nervous system, which promotes "rest and digest" functions.

Beck Depression Inventory A 21-item self-report inventory that measures the severity of depression symptoms.

Beta waves Brainwaves in the 13-30 Hz range, associated with active thinking, focus, and alertness.

Blind Source Separation A statistical technique for separating a set of mixed signals into their original source signals, without prior knowledge of the sources or the mixing process. It aims to recover the underlying sources by exploiting the statistical properties of the mixed signals.

Blood Volume Pulse A measure of the changes in blood volume within a given tissue, often used to assess cardiovascular function. It can be measured using various techniques, such as photoplethysmography (PPG), which uses light to detect changes in blood volume in the microvascular bed of the skin.

Blood-Oxygen-Level-Dependent The contrast mechanism used in functional magnetic resonance imaging (fMRI) to map neural activity in the brain. It is based on the fact that changes in neuronal activity lead to changes in blood flow and oxygenation, which can be detected by MRI scanners.

Boundary Element Method A numerical computational method for solving the forward problem, which involves calculating the electric potential distribution on the scalp given a known source configuration inside the brain. BEM only requires the discretization of the head's surface, making it computationally efficient for EEG analysis.

Brain-Computer Interface A direct communication channel between a brain and an external device, such as a computer. It can be active or passive: the former decodes brain signals that the user intentionally generates, in the latter it analyses a user's state based on external stimuli.

Brief Mood Introspection Scale An open-source mood scale consisting of 16 mood-adjectives to which a person responds.

Canonical Correlation Analysis A multivariate statistical method that analyzes the relationship between two sets of variables. It seeks to identify pairs of linear combinations of variables (one from each set) that have maximum correlation with each other. CCA is used to find patterns of association between two sets of measurements.

Cauchy Principal Value A method for assigning a finite value to certain improper integrals that would otherwise be undefined due to singularities in the integrand. It involves taking the limit of the integral as the integration bounds approach the singularity symmetrically from both sides.

Central Nervous System The complex network of nerve tissues that controls the activities of the body. In vertebrates it comprises the brain and spinal cord.

Common Average Referencing An EEG referencing technique where the average signal across all electrodes is subtracted from each individual electrode's signal. This helps to reduce common noise and artifacts that are present across the scalp, thereby improving the signal-to-noise ratio and highlighting localized brain activity.

Common Spatial Pattern A statistical technique used to analyze multi-channel data, often employed in brain-computer interfaces to extract features from EEG signals. It aims to identify spatial patterns that maximize the variance between two conditions or classes, such as different mental states or motor imagery tasks. CSP computes spatial filters that project the data onto a lower-dimensional subspace where the differences between the classes are most prominent.

Computer Vision A field of artificial intelligence that enables computers to extract meaningful information from visual inputs like images and videos. It uses techniques like machine learning to "understand" the content of these inputs, similar to human vision.

Contingent Negative Variation A slow, negative-going event-related potential (ERP) that occurs in the brain when a person is anticipating a stimulus that requires a response. It reflects the preparation for motor responses and cognitive processing related to the anticipated event.

Convolutional LSTM Neural Network A type of recurrent neural network (RNN) that combines the strengths of convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. It is particularly well-suited for analyzing spatiotemporal data, such as videos, sequential images, and EEG signals.

Convolutional Neural Network Type of artificial neural network specialized for processing grid-like data, such as images. The convolutional layers to extract features from the input data, making them highly effective for tasks like image recognition and object detection.

databaASe for impliCit pERsonaliTy and Affect recognitIoN The first database to connect personality traits and emotional states via physiological responses. ASCERTAIN contains big-five personality scales and emotional self-ratings of 58 users along with synchronously recorded EEG, ECG, GSR and facial activity data, recorded using off-the-shelf sensors while viewing affective movie clips.

Database for Emotion Recognition through EEG and ECG signals A multimodal dataset for emotion recognition research. It contains EEG and ECG signals recorded from subjects while watching film clips designed to elicit different emotional states. DREAMER is notable for its use of low-cost for data acquisition.

Dataset for Emotion Analysis using Physiological signals A multimodal dataset for emotion analysis using physiological and EEG signals. It contains data from subjects while watching music videos, along with their self-reported emotional ratings.

Deep 4-Net A deep learning architecture specifically designed for EEG signal analysis. It is characterized by its use of four convolutional layers with different filter sizes, enabling it to capture both local and global features from the EEG data. Deep4Net has shown promising results in various EEG-based applications, including emotion recognition and motor imagery classification.

Deep Learning Specialized type of Machine Learning that uses artificial neural networks with multiple layers to analyze data, enabling it to learn complex patterns and representations.

Delta waves The slowest brainwaves (0.5-4 Hz), predominantly found in deep sleep stages.

Diffusion Tensor Imaging An MRI technique that measures the diffusion of water molecules in tissues. It provides information about the microstructure of white matter tracts in the brain and can be used to assess brain connectivity and identify abnormalities in white matter integrity.

Digital Imaging and COmmunications in Medicine The international standard for medical images and related information. It defines the formats for medical images that can be exchanged with the data and quality necessary for clinical use.

Discrete Wavelet Transform A wavelet transform that operates on discrete-time signals. It decomposes a signal into a set of wavelet coefficients at different scales and positions, providing a time-scale representation of the signal.

early stage PREProcessing A standardized preprocessing pipeline for EEG data that aims to improve the quality of EEG recordings by addressing common sources of noise and artifacts. It involves steps such as line noise removal, referencing, interpolation of bad channels, and identification of noisy epochs.

EEG Inception Temporal Network A deep learning architecture designed for EEG signal analysis, particularly in brain-computer interface (BCI) applications. It combines inception modules and temporal convolutional networks to effectively extract features from EEG data, achieving high accuracy in tasks like motor imagery classification.

EEGNet A compact convolutional neural network specifically designed for analyzing EEG data. It incorporates specialized layers and filters to effectively capture temporal and spatial patterns in EEG signals, making it suitable for various EEG-based applications, such as brain-computer interfaces and seizure detection.

Effective Connectivity A measure of the causal influence that one brain region exerts over another. It goes beyond simple correlations and aims to identify the direction and strength of interactions between different parts of the brain.

Electrocardiography A non-invasive diagnostic test that records the electrical activity of the heart using electrodes placed on the skin.

Electrocorticography A type of intracranial electroencephalography (iEEG) that uses electrodes placed directly on the surface of the brain to record electrical activity from the cerebral cortex. It offers higher spatial resolution and signal-to-noise ratio compared to scalp EEG and is often used in epilepsy surgery and brain-computer interface research.

Electroencephalography Topographical Representation A visual representation of brain electrical activity recorded by EEG, where the scalp is depicted as a 2D map and colors or contours indicate the amplitude or frequency of the EEG signal at different locations. This allows for the visualization of spatial patterns of brain activity.

Electroencephalography A diagnostic examination that measures and records the electrical activity of the brain.

Electromyography A diagnostic technique that assesses the health of muscles and the nerve cells that control them (motor neurons). It involves measuring the electrical activity of muscles using electrodes placed on the skin or inserted into the muscle. EMG is used to diagnose various neuromuscular disorders.

Electrooculography A technique for measuring the corneo-retinal standing potential that exists between the front and the back of the human eye. It is used to record eye movements and is often employed in sleep research and diagnostics of eye conditions.

Empirical Mode Decomposition A data-driven signal processing technique that decomposes a non-stationary and nonlinear signal into a set of Intrinsic Mode Functions (IMFs). It is an adaptive method that does not require prior knowledge of the signal's characteristics or any predefined basis functions.

Ensemble Empirical Mode Decomposition A noise-assisted data analysis method that overcomes the mode mixing problem of Empirical Mode Decomposition (EMD) by adding white noise to the signal and repeating the decomposition multiple times. The final IMFs are obtained by averaging the IMFs from each ensemble member.

Entropy Rate Minimization A technique used in various fields, such as machine learning, signal processing and information theory, to find models or processes that minimise the entropy (or randomness) of the system over time

Event-Related Desynchronization A decrease in the power of a specific frequency band in the EEG or MEG signal that is time-locked to an event. It is often interpreted as a correlate of increased cortical activity and reflects the active engagement of neuronal populations in information processing.

Event-Related Potential Measured brain response that is the direct result of a specific sensory, cognitive, or motor event. More formally, it is any stereotyped electrophysiological response to a stimulus. The study of the brain in this way provides a noninvasive means of evaluating brain functioning.

Event-Related Spectral Perturbation A measure of how the power spectrum of a signal, such as an EEG or MEG signal, changes over time in response to an event. It is often used to analyze brain activity associated with sensory, cognitive, or motor events.

Event-Related Synchronization An increase in the power of a specific frequency band in the EEG or MEG signal that is time-locked to an event. It is often interpreted as a correlate of decreased cortical excitability or an idling state of neuronal populations.

Event-Stimulus Potential Measured brain response that is the direct result of a specific event related only with physical stimulus characteristics.

Evoked Potential Electrical potential recorded from the nervous system following the presentation of a stimulus, such as light, sound, or touch. Evoked potentials are used to assess the function of sensory pathways and identify neurological abnormalities.

Exact Low Resolution Electromagnetic Tomography A further refinement of LORETA that provides even more accurate source localization. It uses a more sophisticated algorithm to estimate the current density distribution, resulting in improved spatial resolution and reduced localization error.

Extended Infomax Algorithm An extension of the Infomax algorithm for Independent Component Analysis (ICA) that enables the separation of sources with both sub-Gaussian and super-Gaussian distributions. It achieves this by incorporating a flexible nonlinearity in the learning rule, allowing it to adapt to different source distributions.

Eysenck Personality Questionnaire A personality questionnaire designed to measure the major dimensions of personality. It was developed by Hans Jürgen Eysenck and Sybil B. G. Eysenck and assesses personality traits on three dimensions: extraversion-introversion, neuroticism-stability, and psychoticism.

F1-score The harmonic mean of precision and recall. It provides a balanced measure of the model's performance, especially when the classes are imbalanced.

Facial Landmark Trajectories Dynamic movement patterns of specific points on the face (e.g., corners of the mouth, eyebrows) extracted from video recordings. These trajectories are used in emotion recognition systems to analyze facial expressions and infer emotional states.

False Negative The number of cases where the model incorrectly predicts the negative class (a "miss").

False Positive The number of cases where the model incorrectly predicts the positive class (a "false alarm").

Fast Iterative Filtering A signal decomposition technique that separates a signal into its constituent modes based on their characteristic frequencies. It uses an iterative process to extract the modes, starting with the highest frequency mode and progressively extracting lower frequency modes.

Finite Element Method A numerical method for solving the forward problem by dividing the head volume into small elements and approximating the electric potential distribution within each element. FEM can handle complex head geometries and tissue conductivities but is computationally more demanding than BEM.

Finite Impulse Response A type of digital filter whose impulse response (the output in response to a brief input signal) is of finite duration. This means that the output of the filter will settle to zero in a finite amount of time after the input signal has stopped.

FLuid-Attenuated Inversion Recovery A type of magnetic resonance imaging (MRI) sequence that suppresses the signal from cerebrospinal fluid (CSF), allowing for better visualization of brain lesions and abnormalities near the ventricles or subarachnoid space.

Fourier Transform A mathematical transform that decomposes a signal into its constituent frequencies. It represents a signal in the frequency domain, showing the amplitude and phase of each frequency component.

Free Induction Decay The signal produced by the precessing magnetization of nuclei after excitation by a radiofrequency (RF) pulse. The FID signal decays over time due to T2 relaxation and dephasing of the spins.

Functional Connectivity The statistical dependencies or relationships between physiological signals recorded from different brain regions. It reflects the integration and communication between different parts of the brain.

functional Magnetic Resonance Imaging A neuroimaging technique that measures brain activity by detecting changes in blood flow. It is based on the principle that active brain regions require more oxygen and therefore exhibit increased blood flow. fMRI is widely used to study brain function, cognition, and disease.

functional Near-infrared Spectroscopy A non-invasive neuroimaging technique that measures brain activity by monitoring changes in blood oxygenation levels. It uses near-infrared light to penetrate the scalp and skull, and it is relatively portable and inexpensive compared to other neuroimaging modalities.

Galvanic Skin Response A method of measuring the electrical conductance of the skin, which varies with its moisture level. This is of interest because the sweat glands are controlled by the sympathetic nervous system, so skin conductance is used as an indication of psychological or physiological arousal.

Gamma waves The fastest brainwaves (above 30 Hz), linked to higher cognitive functions, learning, and information processing.

Generalized Partial Directed Coherence A frequency-domain measure of directed connectivity between time series, derived from multivariate autoregressive models. It quantifies the causal influence of one signal on another, taking into account the influence of other signals in the system. GPDC is an extension of Partial Directed Coherence (PDC) and is more robust to noise and differences in signal power.

Granger Causality A statistical concept of causality that is based on prediction. It states that if a signal X "Granger-causes" a signal Y, then past values of X should contain information that helps predict future values of Y, above and beyond the information contained in past values of Y alone.

Graph Theory A branch of mathematics that studies graphs, which are mathematical structures used to model pairwise relations between objects. In the context of brain research, graph theory is used to represent and analyze brain networks, where nodes represent brain regions and edges represent the connections between them.

Hertz The unit of frequency in the International System of Units (SI), equal to one cycle per second. It is used to measure the frequency of various phenomena, including brainwaves, sound waves, and electromagnetic waves.

Hilbert Transform A mathematical transform that creates a complex-valued signal from a real-valued signal. It shifts the phase of the signal by 90 degrees, which is useful for analyzing the amplitude envelope and instantaneous frequency of a signal.

Hollywood Human Action Database EEG dataset. Contains 8 classes of human actions from 32 Hollywood movies: AnswerPhone, GetOutCar, HandShake, HugPerson, Kiss, SitDown, SitUp, StandUp.

Human-Computer Interaction The study and design of the interaction between people (users) and computers. It involves the creation of intuitive interfaces and experiences to optimize how users interact with technology.

Human-Robot Interaction The study of interactions between humans and robots. It encompasses a wide range of topics, including robot design, control, perception, communication, and social interaction. The goal of HRI is to create robots that can interact with humans in a natural, intuitive, and effective way.

ICLabel An automated EEG IC classifier that helps researchers label ICs derived from ICA as brain or non-brain sources. It uses a combination of features extracted from the ICs, such as scalp topography, power spectrum, and dipole location, to classify them into different categories, such as brain signals, eye blinks, muscle activity, and heartbeats.

IMFogram The algorithm for visualizing the IMFs computed by FIF and MvFIF.

Independent Component Analysis A computational technique used to separate a multivariate signal into additive subcomponents that are statistically independent from each other. It is often used to uncover hidden factors or sources that underlie a set of observed signals or measurements.

Independent Component The individual subcomponents that are identified by Independent Component Analysis (ICA). These components are statistically independent, meaning they are not correlated with each other and represent distinct sources of information or activity.

Information Object Definition In DICOM, an object-oriented abstract data model used to specify information about real-world objects applicable to the communication of digital medical information. It provides a standardized way to organize and describe medical image data and related information.

Infrared Infrared (IR) is a type of electromagnetic radiation with wavelengths longer than visible light but shorter than microwaves. IR is used in various applications, including thermal imaging, remote controls, and spectroscopy.

Instantaneous Phase The phase of a signal at a particular instant in time. It can be obtained by applying the Hilbert transform to the signal and then calculating the angle of the resulting complex-valued signal. Instantaneous phase is useful for studying the dynamic behavior of signals and their phase relationships.

Inter-Trial Interval The time interval between the end of one trial and the beginning of the next trial in an experiment. It is often used to allow participants time to reset and prepare for the next trial, and it can also be manipulated to study the effects of timing on cognitive processes.

International Affective Picture System A widely used standardized database of images designed to evoke a range of emotions. Each image is rated on dimensions of valence (pleasantness), arousal (intensity), and dominance (control). Researchers use the IAPS to study emotional responses, their impact on cognition, and individual differences in emotional reactivity.

Intracranial Electroencephalography A type of electrophysiological monitoring that uses electrodes placed directly on the brain's surface (ECoG) or implanted within brain tissue (stereo-EEG) to record electrical activity. It provides a more direct and localized measure of brain activity compared to scalp EEG and is used in epilepsy surgery, brain mapping, and research.

Intrinsic Mode Function The oscillatory modes obtained from Empirical Mode Decomposition (EMD). They represent the local oscillations of the signal at different time scales and satisfy two specific criteria: (1) the number of extrema and zero crossings must be equal or differ by at most one, and (2) the mean of the upper and lower envelopes must be zero.

K-fold Cross-Validation A resampling technique used to evaluate machine learning models on a limited data sample. The data is divided into k subsets, and the model is trained k times, each time leaving out one of the subsets as the validation set and using the remaining k-1

subsets for training. The performance measure is then the average of the values computed in the k iterations.

K-means An unsupervised machine learning algorithm used for clustering data points into groups (clusters) based on their similarity. The goal is to partition the data into K clusters, where each data point belongs to the cluster with the nearest mean (centroid).

K-nearest Neighbors A classification method that predicts the class of a data point based on the majority class among its k-nearest neighbors.

Lab Streaming Layer An open-source system developed by SNNC of USDC for synchronizing and collecting real-time data streams from various sources in research experiments. It allows different devices and software applications to share data with precise timestamps, enabling integrated recordings of EEG, motion capture, eye tracking, and other physiological signals.

Late Positive Potential A positive-going event-related potential (ERP) component that typically occurs between 300 and 700 milliseconds after stimulus onset. It is associated with the processing of emotional and motivationally significant stimuli and is thought to reflect sustained attention and cognitive evaluation of these stimuli.

Leap Motion Controller Optical hand tracking modules capturing the movements of the hands with unparalleled accuracy.

Learning Rate A hyperparameter in machine learning that controls the step size at each iteration while moving toward a minimum of a loss function. It determines how much the model's parameters are updated in response to the estimated error each time the model weights are updated. Choosing an appropriate learning rate is crucial for successful training of deep learning models.

Leave-One-Subject-Out Cross-Validation A cross-validation technique where all but one subject are used for training, and the model is tested on the left-out subject. This is repeated for each subject.

Linear Discriminant Analysis A supervised machine learning technique used for dimensionality reduction and classification. It seeks to find linear combinations of features that best separate different classes or groups in a dataset.

Long Short Term Memory Special type of recurrent neural network (RNN) designed to learn long-term dependencies in data, overcoming the limitations of traditional RNNs in remembering information over extended sequences. It use a unique memory cell structure with "gates" that control the flow of information, allowing the network to selectively remember or forget data.

Low Resolution Electromagnetic Tomography A method for localizing the sources of brain activity from EEG or MEG recordings. It uses a distributed source model, assuming that brain activity is generated by a large number of distributed sources, and estimates the current density distribution in the brain that best explains the observed scalp potentials or magnetic fields.

Machine Learning Type of artificial intelligence that allows computers to learn from data without being explicitly programmed, enabling them to make predictions and decisions.

Magnetic Resonance Imaging A medical imaging technique that uses a strong magnetic field and radio waves to create detailed images of the organs and tissues in the body.

Magnetoencephalography A neuroimaging technique that measures magnetic fields produced by electrical activity in the brain. It offers excellent temporal resolution and good spatial resolution, making it suitable for studying a wide range of brain processes.

Major Depressive Disorder A common and serious mental health disorder characterized by persistent sadness, loss of interest or pleasure in activities, and other symptoms that significantly affect an individual's thoughts, feelings, behavior, and physical well-being.

Mann-Whitney U Test A nonparametric statistical test used to compare two independent groups. Assesses whether the probability of an observation from one group being greater than an observation from the other group is different from 0.5. It is often used as an alternative to the independent samples t-test when the data do not meet the assumptions of normality.

Mean Squared Error A common metric used to evaluate the performance of a regression model. It measures the average squared difference between the predicted values and the actual values.

Millon Clinical Multiaxial Inventory A psychological assessment tool intended to provide information on personality traits and psychopathology, including specific psychiatric disorders.

Mismatch Negativity An auditory event-related potential (ERP) component that occurs when the brain detects a change in a repetitive sequence of sounds. It reflects the brain's automatic detection of auditory deviance and is considered a pre-attentive index of auditory sensory memory.

Motor Execution The actual physical performance of a movement, involving the activation of muscles and the generation of overt motor output. In contrast to motor imagery, which is a purely mental simulation of movement, motor execution involves the actual realization of the intended action.

Motor Imagery A cognitive process in which an individual mentally simulates a movement without actually performing it. It involves the activation of brain regions similar to those involved in actual motor execution, and it has applications in various fields, including sports training, rehabilitation, and brain-computer interfaces.

Motor-Related Cortical Potential A type of event-related potential (ERP) that reflects brain activity associated with the preparation and execution of voluntary movements. It is characterized by a series of negative and positive deflections in the EEG signal that precede and accompany movement onset.

Mu Waves Synchronized patterns of electrical activity involving large numbers of neurons, probably of the pyramidal type, in the part of the brain that controls voluntary movement. These

patterns as measured by electroencephalography (EEG), magnetoencephalography (MEG), or electrocorticography (ECoG), repeat at a frequency of 7.5–12.5 (and primarily 9–11) Hz, and are most prominent when the body is physically at rest.

MultiLayer Perceptron A type of artificial neural network (ANN) that consists of multiple layers of interconnected nodes (neurons). It is a feedforward network, meaning that information flows in one direction, from the input layer through the hidden layers to the output layer.

Multimodal Analysis of Human Nonverbal Behavior - Human-Computer Interaction A multimodal dataset for human-computer interaction research. It includes physiological signals (EEG, ECG, GSR, respiration), audio-visual recordings, and behavioral data from participants interacting with a computer system. MAHNOB-HCI is designed to study user emotional states and engagement during human-computer interactions.

Multiple Sclerosis Chronic autoimmune demyelinating disease, which affects the central nervous system causing a wide spectrum of signs and symptoms.

Multivariate Empirical Mode Decomposition An extension of EMD that can decompose multivariate signals, such as those with multiple channels or dimensions. It considers the relationships between different channels during the decomposition process, allowing for the extraction of common oscillatory modes across channels.

Multivariate Fast Iterative Filtering The multivariate version of FIF algorithm.

MUSEC Dataset to investigate the distinction of music preference through familiarity scores, response times (response rates), and brain response (EEG)

N170 Event-Related Potential A negative-going deflection in the event-related potential (ERP) that occurs approximately 170 milliseconds after the onset of a stimulus. It is most prominently elicited by faces and is thought to reflect the structural encoding of faces.

N1 An early negative-going deflection in the ERP waveform, typically peaking around 100 milliseconds after stimulus onset. It is often associated with selective attention and sensory processing of stimuli.

N400 Event-Related Potential A negative-going deflection in the human event-related potential (ERP) that occurs approximately 400 milliseconds after the onset of a stimulus. It is typically associated with the processing of meaningful stimuli, particularly language, and is more pronounced for stimuli that are unexpected or incongruous with the preceding context.

Naive Bayes A probabilistic classifier based on Bayes' theorem with the assumption of independence between features.

Nencki Affective Picture System Basic Emotion A subset of 510 images from the original NAPS set was selected in order to proportionally cover the whole dimensional affective space. Among these, using three available classification methods,the images eliciting distinguishable discrete emotions were identified.

Nencki Affective Picture System A standardized database of pictures designed to elicit a wide range of emotional responses. It includes images with varying levels of valence (pleasantness-unpleasantness) and arousal (intensity), allowing researchers to study emotional processing and its effects on cognition and behavior.

Neuroinformatics An interdisciplinary field that combines neuroscience, computer science, and information technology to develop tools and methods for organizing, analyzing, and sharing neuroscience data. It involves the creation of databases, software applications, and computational models to facilitate the understanding of brain structure, function, and disease.

Node Betweenness Centrality A measure of centrality in a graph that quantifies the number of times a node lies on the shortest path between two other nodes. Nodes with high betweenness centrality can be considered as bridges or bottlenecks in the network.

Non-Rapid Eye Movement A stage of sleep characterized by slower brain waves, reduced muscle activity, and the absence of rapid eye movements. NREM sleep is divided into three stages (N1, N2, and N3), with N3 being the deepest stage of sleep, also known as slow-wave sleep.

Open Hemispheric Brain Dominance Scale A 24-item inventory that assesses hemispheric dominance, or the tendency to favor one side of the brain over the other.

P200 Event-Related Potential A positive-going deflection in the event-related potential (ERP) waveform that typically peaks around 200 milliseconds after stimulus onset. It is thought to be related to higher-order cognitive processes, such as attention allocation, stimulus evaluation, and decision-making.

P2 A positive-going deflection in the ERP waveform that typically peaks around 200 milliseconds after stimulus onset. It is thought to be related to higher-order cognitive processes, such as attention allocation, stimulus evaluation, and decision-making.

P300 Event-Related Potential A positive deflection in the human event-related potential (ERP) that occurs approximately 300 milliseconds after the onset of a rare, task-relevant stimulus. It is often used as a marker of cognitive processes such as attention and decision making.

Parasympathetic Nervous System One of the two main divisions of the autonomic nervous system. The parasympathetic system is responsible for the body's "rest and digest" function. Its general function is to control homeostasis and the body's rest-and-digest response.

Partial Directed Coherence A frequency-domain measure of directed connectivity between time series, based on Granger causality and derived from multivariate autoregressive models. It quantifies the causal influence of one signal on another, taking into account the influence of other signals in the system. PDC is a precursor to Generalized Partial Directed Coherence (GPDC).

Patient Health Questionnaire-9 The Patient Health Questionnaire-9 (PHQ-9) is a short scale specifically for General Medicine and is used for the diagnosis, monitoring and determination of the severity of depression. The PHQ-9 is' suitable for both screening and for case-finding: it can be administered either by medical personnel or trained or self-administered.

Pearson Correlation Coefficient A measure of the linear correlation between two variables. It ranges from -1 to +1, where +1 indicates a perfect positive linear correlation, 0 indicates no linear correlation, and -1 indicates a perfect negative linear correlation.

Peripheral Nervous System The part of the nervous system that lies outside the brain and spinal cord. It consists of nerves and ganglia that connect the central nervous system to the limbs and organs.

Phase Lag Index A measure of functional connectivity between two brain regions that quantifies the consistency of the phase difference between their signals. It is less sensitive to volume conduction effects than traditional coherence measures.

Phase Locking Value A measure of synchrony between two signals based on the consistency of their phase difference. It ranges from 0 to 1, where 1 indicates perfect phase synchrony.

Phase-based (EEG) In EEG analysis, refers to methods that focus on the phase information of EEG signals to investigate synchrony, connectivity, and dynamic relationships between different brain regions. These methods often involve analyzing phase synchronization, phase-locking value (PLV), or phase lag index (PLI) to quantify the consistency of phase differences between signals.

Photoplethysmography A simple and low-cost optical technique that can be used to detect blood volume changes in the microvascular bed of tissue. It is often used non-invasively to make measurements at the skin surface. A PPG signal can be used to calculate heart rate, heart rate variability, and other physiological parameters.

Pittsburgh Sleep Quality Index A self-report questionnaire that assesses sleep quality and disturbances over a 1-month interval.

Pleasure-Arousal-Dominance An alternative terminology for the Valence-Arousal-Dominance (VAD) model, where "pleasure" is used interchangeably with "valence".

Ponto-geniculo-occipital waves PGO waves are phasic field potentials. These waves can be recorded from the pons, the lateral geniculate nucleus (LGN), and the occipital cortex regions of the brain, where these waveforms originate. They are associated with REM sleep and are theorized to be involved in dream generation.

Positive and Negative Affect Schedule-eXpanded An expanded version of the PANAS scale, designed to assess a broader range of positive and negative emotions, is composed by 60 items on 5-likert point scale. It provides a more detailed assessment of affective states, including basic positive and negative affect as well as specific affective states like fear, sadness, guilt, hostility, shyness, fatigue, serenity, surprise, and joviality.

Positron Emission Tomography A nuclear medicine functional imaging technique that produces a three-dimensional image of functional processes in the body. The system detects pairs of gamma rays emitted indirectly by a positron-emitting radionuclide (tracer), which is introduced into the body on a biologically active molecule. Three-dimensional images of tracer concentration within the body are then constructed by computer analysis. In modern

scanners, three dimensional imaging is often accomplished with the aid of a CT X-ray scan performed on the patient during the same session, in the same machine.

Power Spectral Density Describes the distribution of a signal's power over different frequencies. It shows how much power is contained in each frequency component of the signal.

Power-based (EEG) In EEG analysis, refers to methods that primarily utilize the power or amplitude of EEG signals to extract information and quantify brain activity. These methods often involve calculating the power spectral density (PSD) in specific frequency bands (e.g., delta, theta, alpha, beta, gamma) to assess the relative contribution of different brain rhythms.

Precision The proportion of true positive predictions out of all positive predictions made by the model. It measures how well the model avoids false positives.

Premotor Area A region of the frontal cortex located anterior to the primary motor cortex. It plays a crucial role in planning and preparing movements, as well as in higher-level cognitive functions such as decision-making and motor learning.

Primary Motor Cortex A region of the frontal cortex located in the precentral gyrus, responsible for the execution of voluntary movements. It is the final output stage of the cortical motor system, sending signals directly to the muscles to produce movement.

Principal Component Analysis A dimensionality reduction technique that transforms a set of correlated variables into a new set of uncorrelated variables, called principal components. The principal components are ordered by the amount of variance they explain in the original data, with the first principal component explaining the most variance.

Principal Component The new set of uncorrelated variables resulting from a Principal Component Analysis (PCA). They are linear combinations of the original variables, ordered by the amount of variance they explain in the data. The first principal component captures the most variance, the second principal component captures the second most, and so on.

Proton Density An MRI sequence that provides information about the density of hydrogen protons in tissues. It is sensitive to differences in water content and can be used to identify edema, tumors, and other abnormalities.

PsyToolkit A free, open-source online platform for creating and running experiments and surveys in psychology and related fields. It provides a user-friendly interface and tools for designing experiments, collecting data, and analyzing results.

Quadratic Discriminant Analysis A classification technique that models data using quadratic decision boundaries.

Quantitative Physical Rehabilitation An approach to physical rehabilitation that utilizes objective measures and data analysis to assess patient progress, optimize treatment plans, and improve outcomes. It involves the use of sensors, wearable technology, and other tools to quantify movement, strength, and other relevant parameters.

Radial Basis Function A real-valued function whose value depends only on the distance from the origin or a center point. RBFs are commonly used as activation functions in radial basis function neural networks.

Radiofrequency In MRI, refers to electromagnetic waves in the radiofrequency range that are used to excite the nuclei and manipulate their magnetization. RF pulses are used to create different types of MRI sequences, such as T1-weighted, T2-weighted, and FLAIR.

Rapid Eye Movement A stage of sleep characterized by rapid eye movements, muscle atonia, and vivid dreams. Brain activity during REM sleep is similar to that of wakefulness, with increased activity in areas associated with emotion and memory.

Rapid Visual Information Processing Cognitive assessment for measuring the sustained attention: participants are requested to detect target sequences of digits.

Recall The proportion of true positive predictions out of all actual positive instances. It measures how well the model identifies all the positive instances.

Rechtschaffen and Kales Refers to the a landmark publication by Rechtschaffen and Kales in 1968 that established a standardized system for classifying sleep stages based on EEG, EOG, and EMG signals. The sleep stages identified were four NREM stages (1, 2, 3, 4) and one REM stage.

Recurrent Neural Network Type of neural network designed to handle sequential data by having connections that form loops, allowing them to retain information from previous steps in the sequence. This "memory" makes them well-suited for tasks like language processing, speech recognition, and time series analysis.

Reference Electrode Standardization Technique A technique used in EEG analysis to address the problem of the "reference electrode" by transforming the recorded EEG data to a common average reference. This helps to improve the accuracy and reliability of EEG analysis by reducing the influence of the chosen reference electrode on the results.

Relative Band Power The ratio of the power of a specific frequency band in the EEG signal to the total power of the signal across all frequency bands. It is often expressed as a percentage and is used to quantify the relative contribution of different frequency bands to overall brain activity.

REMembrance of EMOtions A new dataset presented in this thesis that contains EEG data from several people that self-induced emotions from their memories.

Russell Circumplex Model A model of affect that represents emotions along two dimensions: valence (pleasure-displeasure) and arousal (activation-deactivation). Emotions are plotted in a circular space, with the angle representing the degree of arousal and the distance from the center representing the intensity of the emotion.

Self-Assessment Manikin A non-verbal pictorial assessment technique that directly measures the pleasure, arousal, and dominance associated with a person's affective reaction to a wide

variety of stimuli. It uses images of a cartoon figure to represent different levels of these dimensions, allowing respondents to quickly and easily rate their emotional responses.

Set of Fear Inducing Pictures A standardized database of images specifically designed to elicit fear and anxiety responses. It is used in research to study the emotional processing of fear, the neural correlates of fear, and the effects of fear on cognition and behavior.

Shapiro-Wilk Test A statistical test used to assess whether a dataset follows a normal distribution. It is one of the most powerful tests for normality and is often used before applying parametric statistical tests that assume normality.

Sigma waves Brainwaves in the 12-16 Hz range, most prominent during stage 2 sleep and associated with sleep spindles.

Signal-to-Noise Ratio A measure that compares the level of a desired signal to the level of background noise. It is defined as the ratio of signal power to noise power, often expressed in decibels (dB). A higher SNR indicates a clearer signal with less noise interference. In EEG, the SNR is often low, meaning that the signal of interest (brain activity) is contaminated by various artifacts (e.g., muscle movements, eye blinks, line noise). Techniques such as filtering aim to increase the SNR by reducing noise and enhancing the features of interest.

Single-Photon Emission Computed Tomography A nuclear medicine tomographic imaging technique using gamma rays. It is very similar to conventional nuclear medicine planar imaging using a gamma camera, but is able to provide true 3D information. This information is typically presented as cross-sectional slices through the patient, but can be freely reformatted or manipulated as required.

Singular Value Decomposition A matrix factorization technique that decomposes a real or complex matrix into three matrices: U, Σ, and V*. It is a powerful tool for dimensionality reduction, data compression, and noise reduction.

SJTU Emotion EEG Dataset An EEG dataset for emotion recognition research collected at Shanghai Jiao Tong University. It contains EEG data from subjects while watching film clips designed to elicit different emotional states.

Slow-Wave Activity The amount of slow-wave activity (delta wave activity) in the EEG during sleep. It is a marker of depth and intensity of sleep and is often used to assess sleep quality and the effects of sleep deprivation or disorders.

Slow-Wave Sleep A stage of deep sleep characterized by high-amplitude, low-frequency brainwaves (delta waves) in the EEG. It is crucial for physical restoration and memory consolidation.

Small-World Index A measure that quantifies the "small-worldness" of a network. It compares the network's clustering coefficient and average path length to those of a random network. Higher SWI values indicate a greater degree of small-worldness, characterized by high clustering and short path lengths.

STAI-State A subscale of the STAI that measures the temporary condition of state anxiety, characterized by feelings of apprehension, tension, nervousness, and worry.

STAI-Trait A subscale of the STAI that measures the relatively stable personality characteristic of trait anxiety, reflecting a general tendency to experience anxiety across various situations.

standardized Low Resolution Electromagnetic Tomography An improved version of LORETA that provides standardized estimates of current density. It addresses some of the limitations of LORETA by standardizing the current density estimates, which improves the accuracy and interpretability of the results.

State-Trait Anxiety Inventory A self-report questionnaire that measures two types of anxiety: state anxiety (temporary condition) and trait anxiety (long-standing personality characteristic).

Steady-State Visually Evoked Potential An oscillatory response in the brain's electrical activity (EEG) that is elicited by a visual stimulus flickering at a specific frequency. When the retina is excited by a visual stimulus oscillating at a certain frequency, the brain generates electrical activity at the same (or multiples of) frequency of the visual stimulus. This response is most prominent in the occipital region of the brain, where visual information is processed.

Stochastic Gradient Descent An iterative optimization algorithm used to find the minimum of a function. It is a stochastic approximation of gradient descent optimization, where instead of computing the gradient of the loss function using the entire dataset, it estimates the gradient using a small randomly selected subset of the data (a mini-batch). This makes it much faster and more computationally efficient, especially for large datasets, and allows for online learning.

Supplementary Motor Area A region of the frontal cortex located on the medial surface of the brain, anterior to the primary motor cortex. It plays a crucial role in planning and coordinating complex movements, particularly those that are internally generated or require sequential execution.

Support Vector Machine A supervised machine learning algorithm that finds an optimal hyperplane to separate data points into different classes. It aims to maximize the margin, or distance, between the hyperplane and the nearest data points from each class.

Sympathetic Nervous System One of the two main divisions of the autonomic nervous system. Its general action is to mobilize the body's nervous system fight-or-flight response. It is, however, constantly active at a basic level to maintain homeostasis.

t-test A statistical test used to compare the means of two groups. It determines whether the difference between the means is statistically significant by considering the variability within each group.

T1 relaxation time In MRI, the time constant that describes the recovery of longitudinal magnetization (aligned with the main magnetic field) after excitation. T1-weighted images provide good anatomical detail and contrast between different tissues.

T2 relaxation time In MRI, the time constant that describes the decay of transverse magnetization (perpendicular to the main magnetic field) due to spin-spin interactions. T2-weighted images are sensitive to changes in water content and are often used to detect edema, inflammation, and lesions.

Theta waves Brainwaves in the 4-8 Hz range, associated with drowsiness, relaxation, and creative inspiration.

Topographical Event-Related Potential A visualization technique that displays the scalp distribution of event-related potentials (ERPs) over time. It creates a series of "snapshots" that show the spatial pattern of electrical activity on the scalp at different time points after the presentation of a stimulus.

Topographical Event-Related Spectral Perturbation A visualization technique that combines the event-related spectral perturbation (ERSP) with topographical maps of the scalp. It displays how the power spectrum of the EEG or MEG signal changes over time in response to an event, while also showing the spatial distribution of these changes across the scalp.

Toronto Alexithymia Scale A 20-item self-report scale that measures alexithymia, a personality construct characterized by difficulties identifying and describing feelings.

Transfer Entropy A measure of directed information transfer between two time series. It quantifies how much the past values of one signal improve the prediction of the future values of another signal, beyond the information provided by the past values of the second signal itself.

Transformers Type of artificial intelligence model that can analyze and understand sequential data, like text or music. They use a mechanism called "self-attention" to process information all at once, rather than in a specific order, which allows them to learn complex relationships and context.

True Negative The number of cases where the model correctly predicts the negative class.

True Positive The number of cases where the model correctly predicts the positive class.

Universal Approximation Theorem Mathematical proof that a feed-forward neural network with a single hidden layer containing a finite number of neurons can approximate any continuous function on a $\mathbb{R}^n$ compact with arbitrary accuracy, provided the activation function is nonlinear.

Valence-Arousal-Dominance A three-dimensional model for describing and measuring emotional states. It adds a dominance dimension to the valence and arousal dimensions, representing the degree of control or influence over a situation.

Valence The intrinsic attractiveness/goodness (positive valence) or averseness/badness (negative valence) of an event, object, or situation.

Virtual Glove A smart virtual glove for the hand telerehabilitation composed by two Leap Motion Controller sensors in orthogonal position.

Virtual Reality A technology that creates immersive, interactive experiences for users by simulating real or imagined environments. It typically involves the use of a head-mounted display (HMD) that provides visual and auditory feedback, and may also include haptic devices for tactile feedback.

Wavelet Transform A mathematical transform that decomposes a signal into a set of wavelets, which are localized waves with varying frequency and scale. It provides a time-frequency representation of the signal, showing how the frequency content changes over time.

Welch's method A method for estimating the power spectral density (PSD) of a signal. It is a non-parametric method that involves dividing the signal into overlapping segments, applying a window function to each segment, and computing the periodogram of each segment. The final PSD estimate is obtained by averaging the periodograms.

Candidate's Publications

[1] G. Placidi, A. Di Matteo, **D. Lozzi**, M. Polsinelli, E. Theodoridou, "Patient–therapist cooperative hand telerehabilitation through a novel framework involving the virtual glove system", *Sensors*, vol. 23, no. 7, pp. 3463, 2023. DOI: 10.3390/s23073463

[2] **D. Lozzi**, F. Mignosi, M. Spezialetti, G. Placidi, M. Polsinelli, "A 4D LSTM network for emotion recognition from the cross-correlation of the power spectral density of EEG signals", *2022 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pp. 652–657, 2022. DOI: 10.1109/WI-IAT55865.2022.00103

[3] **D. Lozzi**, "Rapid Visual Information Processing neuropsychological test for remote cognitive assessment in multiple sclerosis: A preliminary study", *BPA - Applied Psychology Bulletin*, vol. 81, no. 296, pp. 25–34, 2023. DOI: 10.26387/bpa.2023.00005

[4] **D. Lozzi**, F. Mignosi, G. Placidi, M. Polsinelli, "Graph model of phase lag index for connectivity analysis in EEG of emotions", *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 348–353, 2023. DOI: 10.1109/CBMS58004.2023.00242

[5] M. Polsinelli, A. Di Matteo, **D. Lozzi**, E. Mattei, F. Mignosi, L. Nazzicone, V. Stornelli, G. Placidi, "Portable Head-Mounted System for Mobile Forearm Tracking", *Sensors*, vol. 24, no. 7, pp. 2227, 2024. DOI: 10.3390/s24072227

[6] A. Di Matteo, **D. Lozzi**, E. Mattei, F. Mignosi, S. Montagna, M. Polsinelli, G. Placidi, "Calibration of the Double Digital Twin for the Hand Rehabilitation by the Virtual Glove", *2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 521–526, 2024. DOI: 10.1109/CBMS61543.2024.00092

[7] E. Mattei, **D. Lozzi**, A. Di Matteo, M. Polsinelli, C. Manes, F. Mignosi, G. Placidi, "Deep Learning Architecture analysis for EEG-Based BCI Classification under Motor Execution", *2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 549–555, 2024. DOI: 10.1109/CBMS61543.2024.00097

[8] E. Mattei, **D. Lozzi**, A. Di Matteo, A. Cipriani, C. Manes, G. Placidi, "MOVING: A Multi-Modal Dataset of EEG Signals and Virtual Glove Hand Tracking", *Sensors*, vol. 24, no. 16, pp. 5207, 2024. DOI: 10.3390/s24165207

[9] E. Mattei, **D. Lozzi**, A. Di Matteo, C. Manes, F. Mignosi, M. Polsinelli, G. Placidi, "Analysis of the EEG Resting-State Signals for BCI", *9th Graz Brain-Computer Interface Conference 2024*, pp. 355–359, 2024. DOI: 10.3217/978-3-99161-014-4-062

[10] **D. Lozzi**, E. Mattei, R. Ciuffini, A. Di Matteo, A. Marrelli, R. Ornello, M. Polsinelli, C. Rosignoli, S. Sacco, G. Placidi, "The challenge of driving BCI with emotional signals collected by EEG", *9th Graz Brain-Computer Interface Conference 2024*, pp. 366–371, 2024. DOI: 10.3217/978-3-99161-014-4-064

[11] G. D'Aloisio, A. Di Matteo, A. Cipriani, **D. Lozzi**, E. Mattei, G. Zanfardino, A. Di Marco, G. Placidi, "Engineering a Digital Twin for Diagnosis and Treatment of Multiple Sclerosis", in *International Conference on Engineering Digital Twins (EDTconf 2024)*, 2024. DOI: 10.1145/3652620.3688249

[12] A. Di Matteo, **D. Lozzi**, E. Mattei, S. Montagna, A. Cipriani, F. Mignosi, M. Polsinelli, G. Placidi, "Proactive Hand Rehabilitation Through a Double Digital Twin Provided by the Virtual Glove", in *IEEE MetroXRAINE 2024 - Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering*, 2024. DOI: 10.1109/MetroXRAINE62247.2024.10796497.

[13] **D. Lozzi**, A. Di Matteo, E. Mattei, A. Cipriani, P. Caianiello, F. Mignosi, G. Placidi, "ASIS: a smart alarm clock based on deep learning for the safety of night workers", in *IEEE MetroXRAINE 2024 - Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering*, 2024. DOI: 10.1109/MetroXRAINE62247.2024.10796738.

[14] F. D'Amato, A. Cipriani, A. Di Matteo, **D. Lozzi**, E. Mattei, M. Polsinelli, G. Placidi, "A CycleGAN Model to Synthesize Missing and Unpaired MRI Sequences for Under-Represented Multiple Sclerosis Lesions", in *The 18th European Conference on Computer Vision (ECCV 2024)*, 2024. DOI: in press. https://eccv.ecva.net/

[15] C. Centofanti, **D. Lozzi**, C. Cococcetta, A. Marotta, "An Efficient Edge-Based System for Nucleated Oval-Shaped Red Blood Cell Counting" in *IEEE Access*, 2024. DOI: 10.1109/ACCESS.2024.3515835

[16] S. Migliore, M. De Angelis, I. Di Pompeo, **D. Lozzi**, M. Marcaccio and G. Curcio. "Lifestyle and environment influence the psychological well-being of elderly subjects". *Brain Sciences*, 2024. DOI: 10.3390/brainsci14121276

[17] A. Di Matteo, **D. Lozzi**, F. Mignosi, M. Polsinelli, G. Placidi, "A DICOM-based Standard for Quantitative Physical Rehabilitation", in *Computational and Structural Biotechnology* 2025. DOI: 10.1016/j.csbj.2025.01.012

Part VII

Acknowledgement

Acknowledgements

The work presented in this thesis was carried out at the A²VI-Lab (Acquisition, Analysis, Visualization & Imaging Laboratory), Department of Life, Health & Environmental Sciences (MeSVA) under the guidance of Prof. Giuseppe Placidi, and under the guidance of Prof. Filippo Mignosi from the Department of Information Engineering, Computer Science, and Mathematics (DISIM) at the University of L'Aquila. I express my deepest gratitude to Prof. Placidi and Prof. Mignosi for their constant support, expert guidance, and the invaluable teachings they have imparted to me throughout this journey. A special thanks also goes to all my colleagues at the A²VI-Lab supervised by Prof. Placidi, who contributed to creating a stimulating and collaborative work environment. The constant exchange of ideas and their assistance enabled me to grasp the essential computer science fundamentals crucial for this work. Achieving this goal would not have been possible without the support of this remarkable team.
This work was also carried out during my Erasmus+/Exchange period at the Institute of Neural Engineering (Graz University of Technology, Austria) under the guidance of Prof. Selina Christin Wriessnegger. I am extremely grateful to her and the entire team at the Institute of Neural Engineering for their warm welcome and the infinite availability shown to me. During this period, I had the opportunity to deepen the practical part of signal acquisition, thanks to their patience and the investment they made in me, transmitting valuable skills and a more complete vision of my field of study. The cultural exchange experience enriched me both personally and professionally, allowing me to grow and open myself to new perspectives.

I am deeply grateful to the reviewers for their invaluable expertise and constructive criticism, which have played a crucial role in shaping the argumentative framework and enhancing the clarity of my thesis. Their input has resulted in a more precise, clear, and readable document, facilitating a more impactful dissemination of my research outcomes.

