POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Biomedica

Tesi di Laurea Magistrale

Brain Computer Interface and Eye Tracking Systems as Assistive Technologies

Relatore
prof. Valentina Agostini

Candidato
Federica Germoglio
matricola: 240454

Supervisori esterni
Danmarks Tekniske Universitet
prof. Sadasivan Puthusserypady
prof. John Paulin Hansen

Anno Accademico 2018/2019

Ai miei genitori,
per aver sempre creduto in me.

Sommario

I segnali provenienti dal cervello possono essere impiegati per creare un canale d'informazioni diretto tra cervello umano e computer. Riconoscere l'andamento di questi segnali può rappresentare un metodo per esprimere le intenzioni di persone che altrimenti non sarebbero in grado di comunicare. Rilevare la posizione dello sguardo dell'utente sullo schermo del computer è un altro metodo *hands-free* che consente l'interazione uomo-macchina. In questa tesi entrambe le suddette tecnologie sono combinate al fine di sviluppare un sistema ibrido di scrittura per persone affette da gravi disabilità.

Il sistema di *Brain Computer Interface* sfrutta i potenziali evocati uditivi per selezionare lettere al fine di comporre parole. Il sistema aggiuntivo di *Eye Tracking* consente all'utente di cancellare una lettera errata o salvare una parola. All'utente di questo sistema sarà richiesto solamente di ascoltare stimoli uditivi ed effettuare semplici movimenti oculari. Non è necessario nessun altro tipo di movimento.

Gli esperimenti sono stati condotti su 10 soggetti sani, risultando in una sensitività dell'81% per la Brain Computer Interface e in un'accuratezza del 99% per il sistema di Eye Tracking.

Il progetto combina l'acquisizione, l'elaborazione e la classificazione del segnale elettroencefalografico con il rilevamento della posizione dello sguardo dell'utente, con cui il sistema interagisce tramite una interfaccia grafica.

Abstract

Electrical signals from the brain can be employed to create a direct information channel between the human brain and the computer. Recognizing patterns of these signals can let people's intentions be expressed with no need of movements. Tracking the user's gaze on a computer screen is another a hand-free method for human-machine interaction. In this thesis, both the technologies are combined to develop a hybrid spelling system for people affected by severe disabilities.

The Brain Computer Interface system exploits auditory evoked potentials to select letters in order to write words. The Eye Tracking is combined with this system to let the user delete an unwanted letter or to save a word. A user of this system would only listen to auditory stimuli and make simple eyes movements. No other movements are needed.

Experiments have been conducted on 10 healthy subjects, resulting in a 81% sensitivity for the Brain Computer Interface system and a 99% accuracy for the Eye Tracking system.

The project combines electroencephalographic signal acquisition, processing and classification with gaze detection, and makes the interaction with the user possible through a graphic user interface.

Acknowledgements

I would first like to thank my thesis advisors Associate Professor Sadasivan Puthusserypady and Professor John Paulin Hansen at the Technical University of Denmark. The door to Sada's and John's offices was always open whenever I ran into a trouble spot or had a question about my research or writing.

I would like to express my sincere gratitude to my supervisor at Politecnico di Torino, Associate Professor Valentina Agostini, for her enthusiasm when I first told her about this project, and for her support from afar during this period.

A very special gratitude goes out to the Department of Biomedical Engineering, to the Department of Management and to the Acoustic Lab at DTU for giving me the possibility to use their spaces and equipment.

I would also like to thank my colleagues and friends who took part in my experiments. Ana, Dave, Iva, Marko: thanks for coming to the lab on a rainy Sunday. Bayram, Liam, Matteo, Mustafa, Raga, Rasmus: thanks for lending me your time. Carlo, Matteo: thanks for taking part in the experiments even if you were in Copenhagen on holidays. Without your participation, my experiments could not have been successfully conducted.

Un ringraziamento speciale va alla mia famiglia. Mamma, Papà: grazie di esserci sempre, di essere sempre dalla mia parte, di avermi sostenuto in ogni momento di difficoltà in questi mesi, così come in tutto il mio percorso universitario, così come ogni giorno della mia vita. Grazie per aver appoggiato ogni mia scelta, e di avermi dato la possibiltà di costruire il mio futuro. Quando un giorno avrò raggiunto i miei obiettivi, sia personali che professionali, sarà solo grazie voi. Vi voglio bene. Grazie ad Alessia e a Nicolò, perchè siete una parte di me. Anche se siamo lontani, vi sento sempre vicini.

Grazie a tutta la mia famiglia, ai miei amici vicini e lontani, a chi ha mostrato interesse nei miei progetti e a chi ha condiviso con me questo percorso di studi.

Infine Davide. Grazie di essere stato la mia cavia dall'inizio, di esserti sottoposto a un'infinità di esperimenti (anche se alla fine ho dovuto escluderti perchè avevi le prestazioni peggiori del gruppo) e di non avermi mai detto di no quando ti chiedevo di fare un'altra prova. Il resto lo sai già. Ai nostri traguardi.

Torino, 4 Aprile 2019 Federica

Contents

List of Acronyms

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

More than a billion people are estimated to live with some form of disability, (based on 2010 global population estimates) [1] and nearly 5.4 million people are affected by paralysis [2]. The leading cause of paralysis is stroke, affecting 33.7% of people with paralysis, followed by spinal cord injury (27.3%), multiple sclerosis (18.6%), and cerebral palsy (8.3%) [2]. These diseases can lead to Locked-in Syndrome (LIS): a condition in which a patient is aware but cannot move or communicate due to complete paralysis of nearly all voluntary muscles in the body except for eye movements that are maintained. According to [3], one of the highest costs for a 10-year LIS disease duration are for in-home caregivers ($669,150), meaning that the need for technology to improve the patient's self-sufficiency is increasing.

Assistive Technologies (AT) promote a form of independence for people with disabilities, letting them accomplish simple every-day tasks without the presence of a caregiver [4]. In particular, Augmentative and Alternative Communication (AAC) technologies can help patients to share their ideas and feelings without talking.

1.2 Goals of the project

Brain Computer Interface (BCI) and Eye Tracking (ET) are two examples of already existing assistive technologies. BCIs give the possibility to interact with an external device by using control signals generated by electroencephalographic activity. ET systems are able to translate eye movements into effective communication [5], by controlling the gaze direction. Such technologies would improve the quality of patients' life, reducing, at the same time, the cost of intensive care [6].

The goal is to put technology at the service of the patient, by combining BCI and ET to create a hybrid BCI-ET: this system would decipher thoughts and intentions by means both of brain activity and gaze direction, and control an external device through a simple Graphical User Interface (GUI). The overall system must be fast and intuitive to use and let the patient be more independent.

1.3 Hypothesis

Nowadays our interaction with technological devices is a every-day habit. For this reason, research in the field of Human Computer Interaction (HCI) goes towards improving many aspects of the interaction methods [7]. This work aims to follow the purpose of improvement, starting from the following hypothesis:

- A BCI system is able to identify an intention.
- An ET system is able to identify where the user is looking at by tracking the gaze direction. Many ET systems are based on *Dwell time*, i.r. the user has to fix the target for a pre-defined period of time in order to select it. If the period of time is too short, selection will occur unintentionally. If the time is too long, users will get annoyed. It is impossible to define an optimal Dwell time.
- An hybrid BCI-ET system is a combination of the above mentioned technologies that aims at putting together the tasks of each technology to reach an improvement in the overall system.
- Adding a BCI to an ET system could solve the Dwell time problem by providing an additional and independent communication channel. The hybrid system combines the high speed of ET and the high classification accuracy of BCI.
- Patients affected by LIS are not able to move, except for eyes, but they have sufficient cognitive functionalities. These features can be exploited by the hybrid BCI-ET system to partially restore the patient's self-sufficiency, with an hand-free system.

- Patients need to interface with an efficient technology, meaning that the interaction has to be as quick as possible. Furthermore, they should feel comfortable when interfacing with this system, resulting in the user-friendliness aspect to be stressed.

- Technology has to improve patients' lives, not to make them more difficult. In order to avoid to incur the patient's fatigue, physical and mental efforts must be reduced.

1.4 Contribution

The current project aims at combining BCI and ET technologies with a GUI to build a hybrid hand-free spelling system for people in the late LIS stage. The GUI presents the letters and the user selects them - one by one - in order to write words through a BCI. The ET adds two functionalities: deleting an unwanted letter and saving the current word in a .txt file.

1.5 Outline of the thesis

The thesis is outlined as follows. Explanations of each system, such as BCI, ET, and hybrid BCI-ET are discussed in Chapter 2. In this chapter the state of the art of different methods using such systems are also investigated. In Chapter 3, an introduction of the physiological systems involved in this project is presented. Chapter 4 provides an overview of the equipment used for implementing this project and describes how data are acquired and processed. In Chapter 5, experimental results used to verify the performances of the proposed methods are reported. Discussion about experimental results and future work are presented in Chapter 6. The final conclusions about this work are reported in Chapter 7.

Chapter 2

State of the art

This chapter provides an overview on the different types of Brain Computer Interface (BCI) and Eye Tracking (ET) systems, as well as their evolution in the last few years. The chapter is outlined as follows: explanation of types and applications of BCI, ET, and hybrid BCI-ET systems. The description of each system is provided with reference to past to recent related works.

2.1 Brain Computer Interface

BCIs use brain signals to retrieve information on user intentions. There are many different classification methods for BCI systems. Table 2.1 shows the main features of each neuroimaging method for BCI. First of all, BCI systems can be split into *invasive* and *non-invasive*.

	Neuroimaging Method	**Activity**	**Type**	**Portability**
Invasive	Electrocorticography (ECoG)	Electrical	Direct	Portable
	Intracortical neuron activity	Electrical	Direct	Portable
Non-Invasive	functional Magnetic Resonance Imaging (fMRI)	Metabolic	Indirect	Non-Portable
	Near Infrared Spectroscopy (NIRS)	Metabolic	Indirect	Portable
	Magnetoencephalography (MEG)	Magnetic	Direct	Non-Portable
	Electroencephalography (EEG)	Electrical	Direct	Portable

Table 2.1: Summary of neuroimaging methods for BCI. Adapted from [6].

2.1.1 Invasive BCI

Invasive BCIs record the signal directly from the the brain, needing the electrodes to be placed inside the scalp by surgery. The advantage of invasive methods is that they can record the signal with higher spacial and temporal resolution, as the electrodes are closer to the source of the signal, so that the attenuation due to interposed tissues and the artefacts due to eye blinking and movement are lower. Invasive techniques include:

- *ECoG*: this technique measures electrical activity in the cerebral cortex. It requires a craniotomy to implant an electrode grid, representing a significant risk for the subject. First studies on ECoG have been conducted on animals, evaluating a long term stability of the subdural and epidural signal over several months [8][9][10][11][12]. Experiments with monkeys have developed a less invasive protocols for electrodes implantation [13] as well as showing good performances of the acquisition [9]. ECoG has also been recorded from canines for the study of epilepsy with no results regarding long-term signal quality [14][15].

  In humans, ECoG has been used to investigate motor movements, auditory and visual ability, language function. [16][17] demonstrated that it is possible to decode the direction of hand movements using signals from ECoG. [18] demonstrated that it is possible to infer vowels and consonant in both real and imagined speech. Other studies showed that ECoG signals in gamma range are related to auditory [19][20][21] and language ability [22][23].

  With regard to the use of ECoG in BCI, it has been demonstrated to be a potential tool due to signal quality, resistance to artifacts, spatial and temporal resolution [24][25]. Recent studies showed promising results in motor BCIs [26][27] using sensorimotor ECoG prosthetic arm [28]. It has also been demonstrated that an ECoG BCI system can record motor related gamma signals over years, even if with high inter-subject variability [29].

- *Intracortical neuron activity*: this technique measures the electrical activity in the grey matter of the brain. It needs the electrodes to be implanted inside the cortex to record signals directly from neurons. Even if spatial and temporal resolution of this signal is high [30], signal quality could be effected by the reaction of the tissues to the implanted electrodes [31]. First studies on intracortical neuron activity have been performed on monkeys and rats during learned movements: they showed that this signal can indicate the nature and the direction of the movements [32][33][34]. The same results have been gained by experiments performed on animals making real movements [35].

  More recently, intracortical recordings have been used as part of BCIs to control external prosthetics [36][37][38] able to restore independence of long-term paralysed patients [39][40][41]. Nevertheless a long-term application of this technology could result in the formation of astroglial tissue from the tissue reaction to the presence of the electrode grids [42][43].

2.1.2 Non-invasive BCI

Non-invasive BCIs record the signal from the head, with no need of any surgical implant. These technologies are non-invasive and can extract information from either *indirect* or *direct* methods. Indirect methods measure the *hemodynamic response*, identifying active neurons by the change in ratio of *oxyhaemoglobin* to *deoxyhaemoglobin*, not directly from neuron activity. Indirect techniques include:

- *fMRI*: this technique has been used by a large number of studies since 1900 because of its availability (it can be performed on a clinical magnetic resonance scanner), non-invasiveness and good spatial resolution. It has been used for multiple purposes: disease biomarker [44], therapy monitoring [45], pharmacological studies [46]. fMRI was demonstrated first in rats [47][48] and then in humans [49][50].

  fMRI detects localized brain activity using the Blood Oxygen Level Dependent (BOLD) signal. When neuronal activity increases, the demand for oxygen increases as well and leads to the local blood flow to increase. Blood contains haemoglobin, that is *diamagnetic* when oxygenated and *paramagnetic* when deoxygenated. This difference in magnetic properties leads to differences in the magnetic resonance (MR) signal that are used by fMRI. Indeed BOLD signal increases proportionally to neural activity and reaches a plateau if the stimulus is maintained for sufficient time.

  fMRI has been used for BCI because researches [51] reported that a subject can learn to increase or decrease the BOLD response as feedback. Recently, the possibility of using real-time fMRI lead to multiple application in BCI. An fMRI-based BCI has been developed to let a subject select letters from a keyboard, by activating hands and toes [52]. fMRI has been also used for the development of neurofeedback training for rehabilitation [53][54][55][56].

- *NIRS*: it is a spectroscopy method that employs infrared light to detect cerebral metabolism during neural activity, limited to the outer cortical layer because of the superficial penetration of the light in the brain. NIRS identifies alteration in oxyhaemoglobin and deoxyhaemoglobin concentration by measuring the light attenuation. NIRS can be used for BCI employing the *neurovascular coupling* phenomenon, so that the subject can induce a vascular response when performing a cognitive task.

NIRS-BCI systems use motor imagery [57][58] and cognitive tasks (mental arithmetic [59][60], mental singing [59][60], n-back task [61], as well as many mental tasks in the same trial [62]).

Direct techniques include:

- *MEG*: it employs magnetic induction to record the magnetic activity of the brain. In particular, MEG measures the magnetic field produced by the intracellular current flow [63].

  The first online MEG-based BCI has been presented in 2005 [64], followed by other studies [65][66][67]. Although [68] concludes that MEG-based BCIs are a powerful tool to explore brain functions, they remain still early stage because of costs and lack of practicality.

- *EEG*: it uses electrodes placed on the scalp to register the electrical activity of the brain. EEG is the expression of the synaptic processes - mainly post-synaptic electric potentials - that flow perpendicularly with respect to the scalp. The electrical activity of the brain includes brainwaves that can be classified according to their frequency range. *Delta waves* (0.1-3 Hz) are related to deep sleeping state. *Theta waves* (4-7 Hz) are related to REM phase. *Alpha waves* (8-12 Hz) are typical of waking state with eyes closed and of the pre-sleeping phase. *Mu waves* (7-13 Hz) are synchronized patterns of electrical activity that involve many neurons in the motor cortex. *Beta waves* (13-30 Hz) are registered in the waking state and during an intense mental activity. *Gamma waves* ($>$ 30 Hz) are related to high tension states.

  EEG-based BCIs are the most diffused technique, because of the widespread availability, the non-invasiveness and the good performances. This technique has been used since 1973 [69] and has many different applications.

  An important issue in non-invasive BCI research is to make the EEG acquisition more comfortable and suitable for every-day use, by minimizing the number of electrodes to reduce hassle and setup time.

2.1.3 Control signals for BCI

BCIs use signals - called *control signals* - recorded from the brain to decipher the user's intentions. Control signals can be classified into two categories: *exogenous signals* and *endogenous signals*, as shown in Table 2.2.

	Signal	**Physiological phenomena**	**Training**
Evoked	Visual Evoked Potentials (VEP)	Brain signal modulation in the visual cortex	No
	P300 Evoked Potentials	Positive peaks due to an infrequent stimulus	No
Spontaneous	Slow Cortical Potentials (SCP)	Slow voltage shifts in the brain signals	Yes
	Sensory Motor Rhythms (SMR)	Modulation in sensorimotor rhythms due to motor activities	Yes

Table 2.2: Control signals for BCI. Adapted from [6].

Esogenous signals are due to the response of the brain to an external input, that is receipted by the sensory system. They include:

- *VEP*: brain signals modulations that occur in the visual cortex after a visual stimulus. There are two main types of VEP, depending on the frequency of the stimulus: a less frequent stimulus (< 6 Hz) provokes Transient Visual Evoked Potentials (TVEP) that disappear before the next stimulation. A more frequent stimulus (> 6 Hz) induces Steady-State Visual Evoke Potentials (SSVEP) [70]: in this case the responses are overlapped. SSVEP are largely more used as control signals for BCI because their frequency and phase are more stable than TVEP throughout the whole experiment duration, and they are less exposed to artifacts and noise [71].

- *P300 Evoked Potentials*: positive peaks in the EEG that occur approximately 300 ms after an infrequent external stimulus (Figure 2.1). P300 is related to the *"oddball paradigm"*, a technique based on event-related potential (ERP)s generated in response to unpredictable but recognizable events that are presented to the subject. The reaction can be due to either visual, sensual or auditory stimuli consisting on target and non-target stimuli. The subject has to focus attention on the target event, in order to evoke the neural reaction. The amplitude of P300 is related to the relevance of the stimulus and to its probability: [73] demonstrated that the most probable the stimulus, the smallest the amplitude of the evoked signal. The variability of the signal due to the neural reaction is

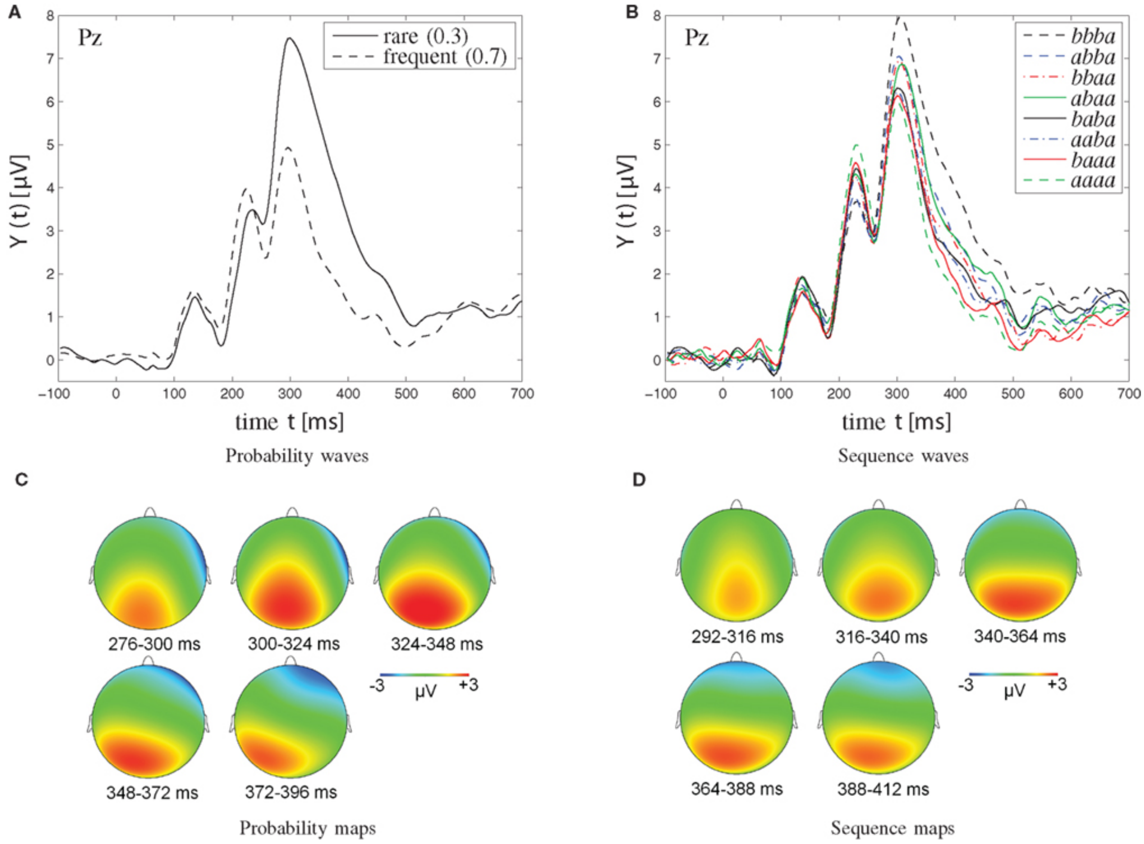


Figure 2.1: Waveforms (A,B) and topographic maps (C,D) of P300 amplitudes. Reprinted from [72]

also related to age and subject health [74]. P300 is generally observed in the parietal lobe and it attenuates in amplitude when increasing distance. Although the neural mechanism of P300 are still unclear, this control signal is common in BCI applications.

Endogenous signals are independent from any external stimulation: they can be operated at free will, after a training period. They include:

- *SCP*: current shifts in the cortical activity that can last from hundreds of milliseconds to seconds. Negative SCP are caused by synchronous slow excitatory postsynaptic potentials and are related to neuron excitation, whereas positive SCP indicate an inhibition with a reduction in cortical activity. It is possible to train people to modulate SCP in order to generate voluntary changes in the signal. SCP can be used for self-regulation training in pathological conditions [75] is such a way that the user can express a neurofeedback. In this case BCI is referred to as a Thought Translation Device (TTD).

The first study on EEG-based BCIs using SCP was performed 20 years ago [76] with two patients in LIS. Later on, many TTD devices have been developed [77][78][79], to let paralyzed people select letters, words or pictures, or move a cursor on a computer screen. A huge drawback of the SCP technique is that it requires a long-term preparation and extensive training to gain the control of the signal [80].

- *SMR*: oscillations registered over the sensorimotor cortex in the mu, beta and gamma frequency bands [81]. SMR can undergo two types of modulation, i.e. event-related desynchronization (ERD) and event-related synchronization (ERS). ERD is a reduction in the oscillatory activity related to a sensorimotor event [82]. An increase in the oscillatory activity leads to ERS, that can be related to a dis-inhibition of cortical network [83]. Low frequency ERD (8-10 Hz) in the mu band are related to motor behaviour and can be recorded over the sensorimotor cortex. Instead, high frequency ERD (10-13 Hz) are related to a specific task and their recording is restricted. ERD in the beta band occurs after a movement, even if only imagined.

  Many studies have been conducted about SMR-based BCIs for motor imagery (MI) applications: in this case the subject's motor intention is translated into a control signal, without any external stimulus. The most common mental task investigated for MI-based BCI are limb movements. A common application for this technique are MI-based spellers. This interfaces give the user the possibility to select a single letter or character, a group of letters, or a command, on a screen. [84][85][86][87]. Another application for MI-based BCIs is the control of external devices, e.g. exoskeleton [88][89][90][90], robotic limbs [91], wheelchairs [92][93][94]. Even if SMR-based BCIs require both an intensive training to let the user learn to modulate the signal, and multiple EEG channels for good recording performances, they are a suitable technique for control applications because they are completely independent from external stimuli (they let the user operate at free will) and from the movement ability (they can be used by LIS patients).

2.2 Eye Tracking

ET systems exploits the communicative power of the eyes for human-computer interaction. Real-time gaze-based interfaces are a powerful mean of communication and control for people with physical disabilities. ET systems fulfill a quick and easy interaction, without requiring neither training nor effort to the user.

2.2.1 Types of ET

There are four different types of eye-trackers [95]:

- Sensor: a safe way to apply sensors to the eye is using contact lenses. They can track the gaze either by a mirror that reflects light or a coil of wire orientation in a magnetic field. The advantage of such a technique is high accuracy and temporal resolution. The drawback is that it is an invasive method, as it requires the lenses to be put in the eyes.
- Electro-oculography (EOG): in this case, sensors are attached to the skin around the eyes in order to measure the electric field produced by the eyes rotation. This technique is not suited for everyday use because it requires the electrodes to be carefully placed. The advantage is that with this method it is possible to detect movements even if the eyes are closed, e.g. while sleeping. This technique has been used for emotion recognition and classification [96]
- Photo-oculography (POG): it measures the intensity of reflected infrared light that illuminates the eye. The eye position changes are detected by measuring the difference between the incident and the reflected light. It is an invasive technology because the light source and the sensors have to be put onto glasses. Infrared oculography in less noisy with respect to EOG but it is more sensitive to external light changes.
- Video-oculography (VOG): it is the most common technique for ET. Video-based ET systems can be either stationary or mobile and they can use either visible light or infrared light. In *stationary* ET systems, the eye tracker is positioned near the object to be tracked (usually a screen) and the user is placed in a stationary position in front of the screen. In *mobile* ET system, the user is able to move around because the eye movements are tracked by ET glasses. This setup is also called *head-mounted*.

2.3 Hybrid BCI-ET systems

BCI and ET are technologies that give the user the possibility to interact with an external interface. They can be combined together to construct a hybrid BCI-ET system, that takes advantages from both.

ET has been combined with a MI-based BCI to improve a word selector [97]. In this study the *Dwell time*-based selection (when the gaze fixes the target for a certain time) is substituted by the detection of both gaze fixation and MI in order to increase the resting time for the eyes. The same technique is useful to increase the accuracy in a 3D selection environment [98], if the imaginary movement selects the third axis.

Another study used two different mental states detected by EEG together with ET to control a telepresence robot [99]. In this case, the user can explore the environment with spontaneous gaze and trigger the motion with MI.

In [100], a hybrid interface has been developed by combining eye movements and brain mental activity to allow real-time control of a quadcopter in a 3D physical space, achieving compatible performances to standard keyboard-based systems.

ET and SSVEP have been combined to improve both speed and accuracy of a speller. The goal is to prevent errors with a double-check system: the letter is selected if information from ET and SSVEP-based BCI match. The same technique has been used for a system that utilizes eye tracking for initial rough selection and the SSVEP technology for fine target activation, resulting in a speedy and user-friendly interface [101].

The control of home appliances is obtained by combining a SSVEP-based BCI, a SMR-based BCI and ET, for selecting the device and the commands related to it [102]. As the different modalities can be activated independently, the hybrid system minimizes the number of involuntary selections.

BCI and ET systems are developed to work with signals recorded respectively from brain and eyes. For that reason it is important to understand the structure and the physiological functions of nervous and visual systems, that will be explained in the next chapter.

Chapter 3

Physiological background

The design of a hybrid BCI-ET system involves choices concerning the anatomy and the physiology of the related systems in the human body. This chapter provides an overview of nervous and visual systems organization and functions.

3.1 Nervous system

The nervous system can be divided into the central nervous system (CNS) and the periferal nervous system (PNS) (Figure 3.1). The CNS consists of the brain and the spinal cord. The PNS consists of afferent (or sensitive) neurons and efferent neurons. The efferent neurons are divided into the somatic nervous system (neurons connected to skeletal muscles, skin, and sense organs) and the autonomic nervous system (neurons that control smooth muscles and visceral functions) [103].

3.1.1 Neurons

The neuron is the functional unit of the nervous system. It is a specific cell able to receive, integrate and transmit nervous spikes. A neuron consists of a body, called *soma*, connected to tree-like structures with branches, called *dendrites*, and a single long branch, called *axon* [103]. The soma contains the nucleus of the cell and all the cellular organelles. Dendrites receive signals from afferent neurons and propagate them toward the centre of the neuron. The axon conducts the signal from the neuron to the other cells. Many axons are covered by myelin sheaths, leaving exposed sections (nodes of Ranvier) between segments of myelin. Myelin is a lipid-rich substance that insulate

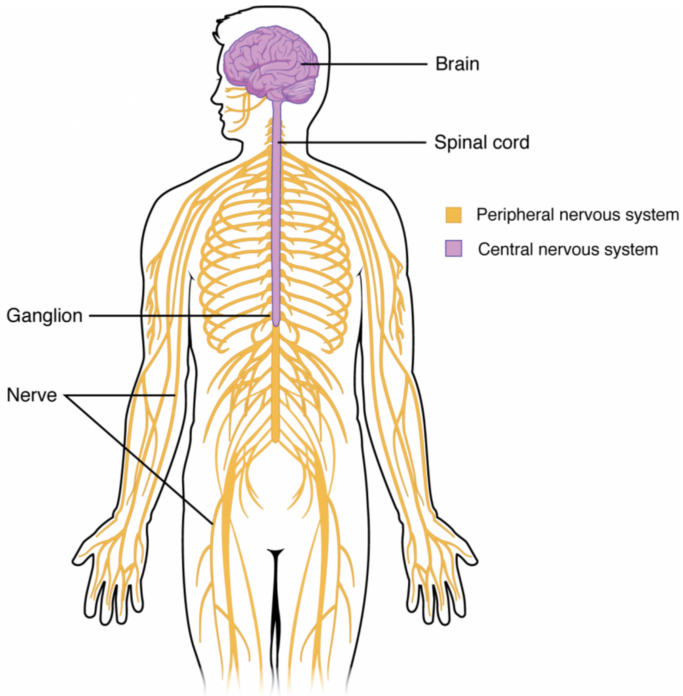


Figure 3.1: Structure of the nervous system. Reprinted from [104].

them to increase the speed of the spike over long distances [105]. Once information reaches the terminal end of the neuron, it is transferred to another cell. The site of communication between a neuron and its target cell is called *synapse* (Figure 3.2).

3.1.2 Action potential generation and transmission

Communication between neurons depends on the action potential propagation. The nervous tissue is an excitable tissue, so it can generate and conduct spikes. The steady state of the cell is a dynamic process that is balanced by ion leakage and ion pumping. The resting membrane potential does not change (-70 mV) without an external influence. To get an electrical signal started, the membrane potential needs to change [104]. A cell is excited when a stimulus makes Na^+ channels open so that sodium ions can enter the cell driven by the concentration gradient. This process increases the value of the membrane potential (*depolarization*). If the depolarization is strong enough to make the membrane potential reach a certain threshold (-55 mV), voltage-gated sodium channels open in the membrane. This results in more Na^+ ions entering the cell, making the membrane potential reach a peak of

+30 mV (Figure 3.3). As the membrane potential reaches the peak, voltage-gated potassium channels open in the membrane while voltage-gated sodium channels close. K^+ ions start to leave the cell and the membrane potential begins to move back toward its resting voltage (*repolarization*).

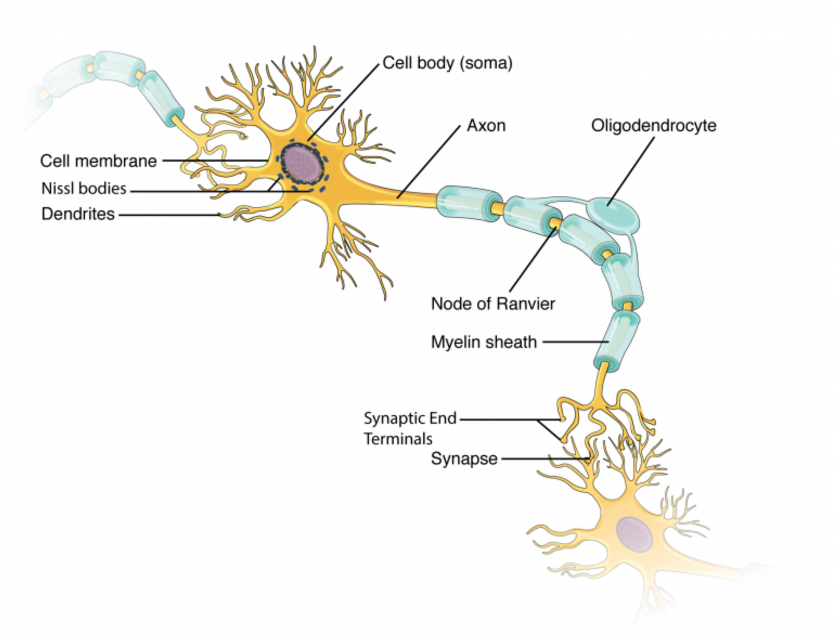


Figure 3.2: Anatomy of the neuron. Reprinted from [104].

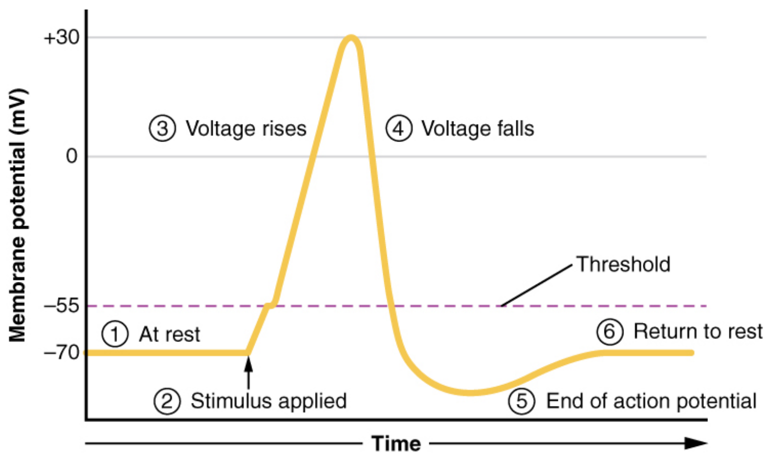


Figure 3.3: Stages of an action potential. Reprinted from [104].

The action potential propagates toward the axon terminal, resulting in *continuous conduction* for unmyelinated axons (where voltage-gate channels are present throughout the membrane) and *saltatory conduction* for myelinated ones (where voltage-gate channels are only found at the nodes of Ranvier). Saltatory conduction is faster than continuous conduction. When the action potential reaches the end of the axon, it is transmitted to the near neuron through synapses. Synapses are the contacts between neurons, and can be either electrical or chemical. Chemical synapses are the most common (Figure 3.4): they work with a neurotransmitter that is released from the presynaptic element, diffuse across the synaptic cleft, binds to a receptor protein and causes a change in the postsynaptic membrane [104].

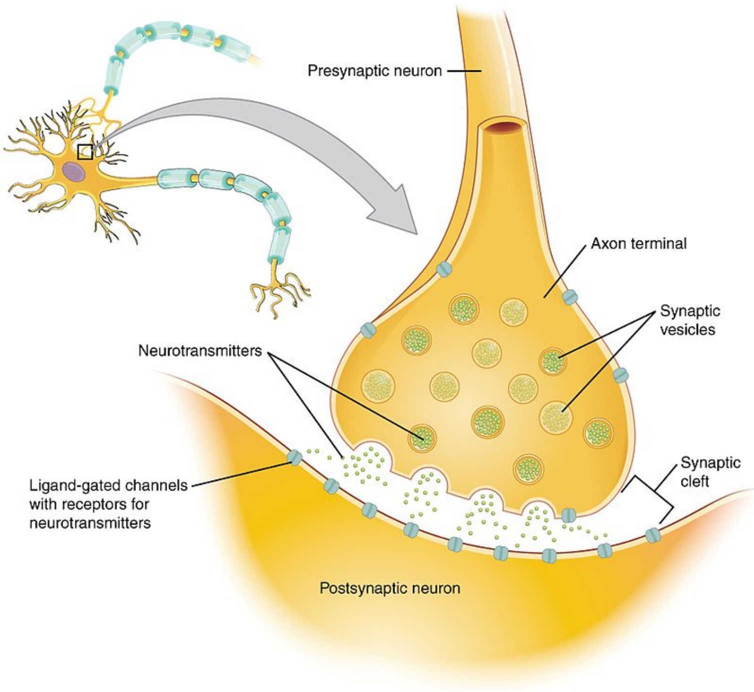


Figure 3.4: Synapse. Reprinted from [104].

3.1.3 The brain

The brain and the spinal cord represent the main organs of the nervous system. The spinal cord is a single structure, while the brain is composed by four major regions (Figure 3.5): the cerebrum, the diencephalon, the brain stem, and the cerebellum [104].

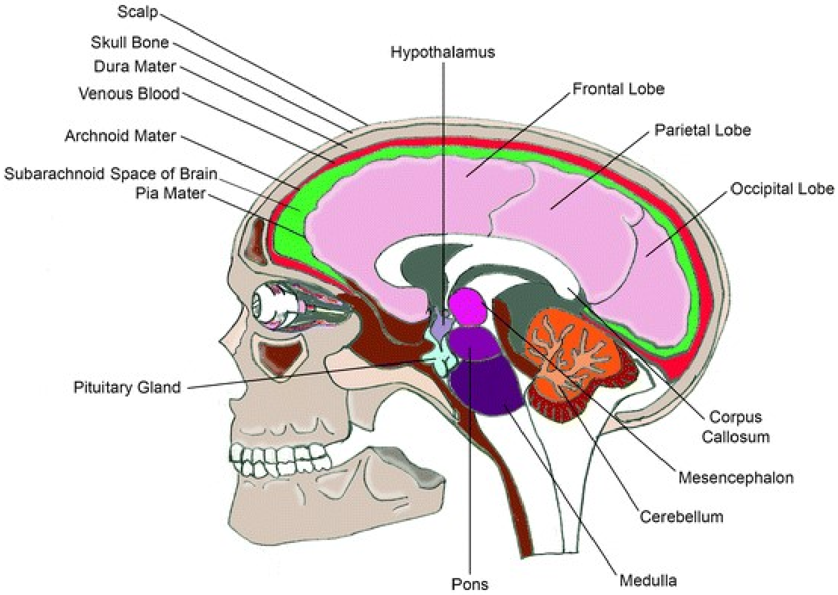


Figure 3.5: Parts of the brain. Reprinted from [106].

The *cerebrum* represents the biggest part of the brain and it is covered by a continuous layer of gray matter, the *cerebral cortex*. The *longitudinal fissure* separates the cerebrum in two distinct sides, a right and left *cerebral hemisphere*. Each cerebral hemisphere includes four lobes (Figure 3.5): *frontal*, *parietal*, *temporal* and *occipital*.

The *frontal lobe* is located in the front of the brain. It is responsible for reasoning, movement, cognition and expressive language. The *parietal lobe* is located in the middle part of the brain. It is responsible for tactile sensory information, proprioception and language processing. It contains the somatosensory cortex. The *temporal lobe* is located in the bottom part of the brain. It is responsible for sound interpretation, language comprehension and memory. It contains the hippocampus and the primary auditory cortex. The *occipital lobe* is located in the back part of the brain. it is responsible for visual stimuli interpretation. it contains the primary visual cortex [104].

3.2 Visual system

The visual system includes three main parts: eyes, optical nerve and visual cortex. It is based on the transduction of a light stimulus received through the eyes [104].

3.2.1 The eye

The eye is the external organ of the visual system. It takes information from the external environment through light (Figure 3.6). The light enters the eye through the *pupil*, and its intensity is regulated by the *iris*, that constricts the pupil in response to bright light, and dilatates it in response of dim light. A system of lenses is able to focus the light and make it converge on the *retina*. When interacting with a photon, the retina undergoes chemical changes provoked by a process called *photoisomerization*. This shape change initiates visual transduction in the retina. The incident light is converted into electrical impulses by photoreceptors (*rod* and *cons*) placed on the surface of the retina. The optical nerve sends these impulses to the brain. Colour vision is provided by different types of *opsin*, a protein that is sensitive to different wavelengths of light [104].

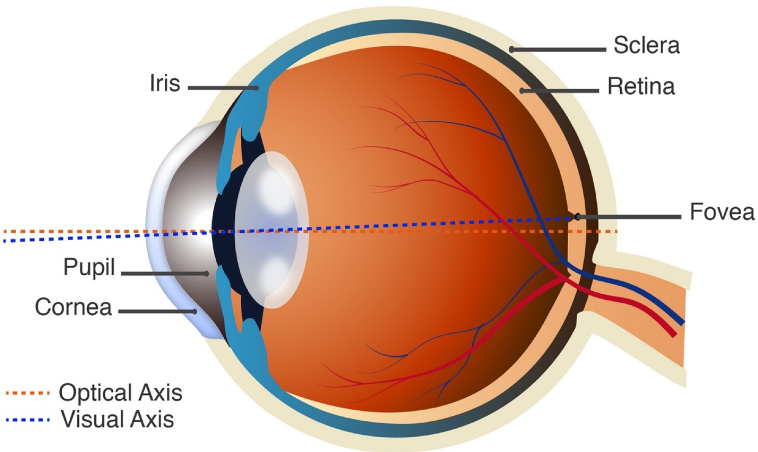


Figure 3.6: Eye anatomy. Reprinted from [107].

3.2.2 Vision

The photoreceptors are concentrated in a small portion of the retina, called *fovea*. As a consequence, a clear vision is possible only in this area. This is the reason why eyes need to move to a position that projects the target directly on the fovea. The control of eye movements involves six muscles and three nerves (Figure 3.7), to move the eyes in coordination. There are two types of movement:

- *Compensation movement*: it is a stabilization movement that occurs either when the head moves but the gaze is kept on a target or when watching a moving object.

- *Saccade*: it is the quick movement (700°/s) that the eyes do when passing from a target to another. It is the most common eye movement and it is intercepted with resting periods (0-1000 ms) called *fixation* [7].

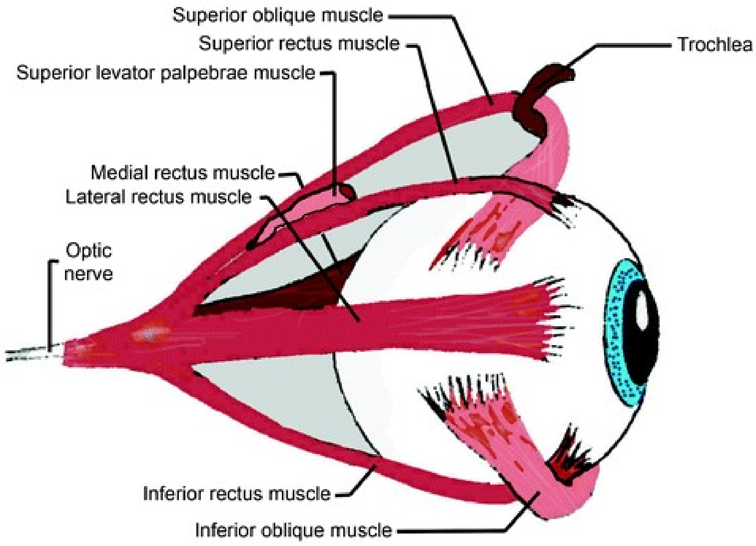


Figure 3.7: Position of extraocular muscles. Reprinted from [106].

Chapter 4

Materials and Methods

This chapter provides an overview of the equipment used for implementing the project and describes how data are acquired and processed.

The overall system includes a BCI part and a ET part, that have been developed and tested separately, and combined with the final interface on a Surface Pro 4 (Microsoft Corporation) with Microsoft Windows 10 Pro operating system.

4.1 BCI system

The spelling system is based on a non-invasive BCI that retrieves Auditory Evoked Potentials (AEP) to establish a communication between user and computer.

4.1.1 Participants

Ten healthy subjects (7 males, 3 females, age 26.61±3.03 years) participated in the experiments. All subjects were volunteering students or researchers. No one had previous experiences with auditory BCI.

4.1.2 Equipment

The BCI system includes an EEG acquisition system connected to the computer as shown in Figure 4.1.
The EEG acquisition system (g.tec medical engineering GmbH, Austria) includes the following components [108] (Figure 4.2):

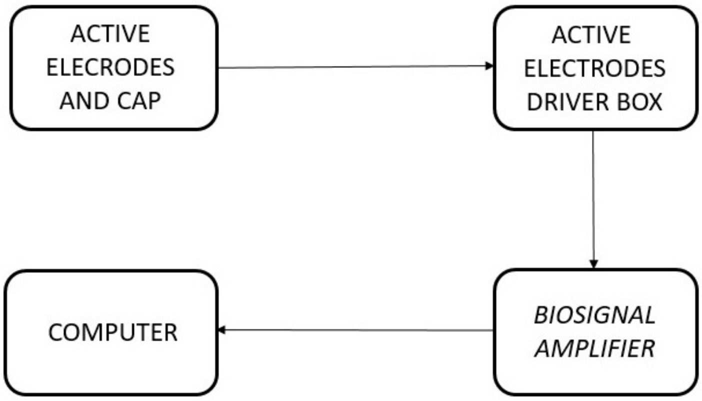


Figure 4.1: Structure of the BCI system.

- *g®.USBamp*: high-performance and high-accuracy biosignal amplifier. It is USB enabled and comes with 16 simultaneously sampled biosignal channels with 24 bits. A total of 4 independent grounds guarantee no interference between the recorded signals. A synchronization cable guarantees that all devices are sampled with exactly the same frequency.

- *g®.GAMMAbox*: power supply and driver box for 16 active electrodes with connectors for g®.USBamp.

- *g®.GAMMAcap*: cap for electrodes placements with 74 labelled standard positions (based on the extended 10-20 system / 10-10 system) and 86 additional intermediate positions. g®.GAMMAcap was available in three sizes (small: head circumference 50 - 54 cm, medium: 54 - 58 cm, large: 58 - 62 cm), and chosen according to the subject.

- *g®.LADYbird*: 16 active ring electrodes, placed on the g®.GAMMAcap, sintered Ag/AgCl crown, red.

- *g®.LADYbirdGND*: passive ground ring electrode, placed g®.GAMMAcap (EEG), sintered Ag/AgCl crown, yellow.

- *g®.GAMMAearclip*: active earclip Ag/AgCl electrode (reference), sintered Ag/AgCl crown, blue.

- *SIGNAGEL®*: highly conductive electrode gel (Parker Laboratories, Inc.).

(a) (b)

(c) (d)

(e) (f) (g)

Figure 4.2: EEG acquisition system (g.tec medical engineering GmbH, Austria). (a) g®.USBamp. (b)(c) g®.GAMMAbox and connectors. (d) g®.GAMMAcap (e) g®.LADYbird (f) g®.LADYbirdGND (g). g®.GAMMAearclip. Reprinted from [108].

4.1.3 Experimental procedure

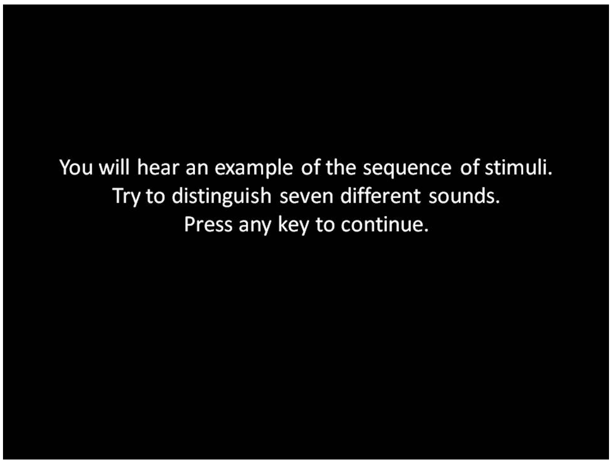
The BCI system aims at understanding the user's intentions through an unexpected auditory stimulus, that evokes a P300 in the EEG signal. The goal is to select a specific stimulus among seven different sounds, retrieving the *oddball paradigm* (see Section 2.1.3).

The overall procedure includes seven trials, one for each stimulus. During each trial, the user has to concentrate on a specific sound, shown on the screen before starting the trial. At the beginning, the user can hear a sequence of stimuli, in order to get used to them. Each trial starts with a interface that shows the target sound and plays it five times, to let the user memorize it. At this point, the user is asked to close his/her eyes, concentrate on the target sound, trying to count how many time he/she can hear it. A sequence of seven sounds is played, with each sound repeated twenty times, in random order. At the end of the sequence, the following trial starts, with another target sound. In Figure 4.3 there are examples of the interfaces shown during the experiment.

The stimuli are represented by seven environmental sounds. This type of sounds are preferred over simple tones, because the user can distinguish them more easily. The sounds employed in this experiment are shown in Table 4.1.

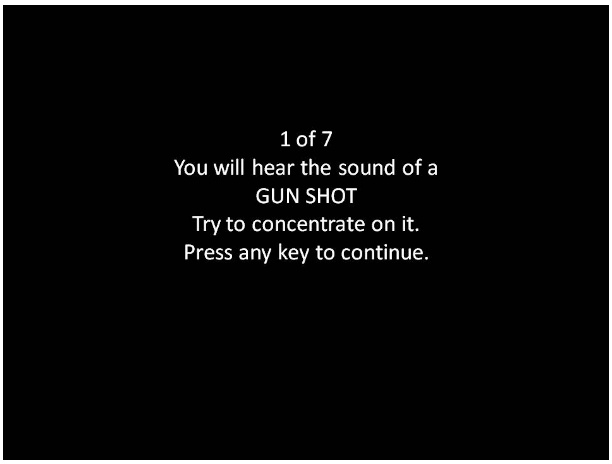
Stimulus	**Sound**
1	Gun shot
2	Church bell
3	Truck horn
4	Doorbell
5	Whistle
6	Phone
7	Moving water

Table 4.1: Sounds employed in the experiment.



(a) Interface 1: beginning of the experiment.

(b) Interface 2: end of the introduction.



(c) Interface 3: beginning of the trial.

Figure 4.3: Example of interfaces shown during the experiment.

The duration of each sound is 200 ms and the inter-stimuli pause is 200 ms. Each sequence of stimuli includes 140 sounds - 20 targets and 120 non-targets - with an overall duration of 56 seconds (28 seconds for sounds and 28 seconds for pause). The experiment, composed by seven trials, includes 392 seconds for the sequences with the addition of the introduction phases (about 35 seconds) and the time needed by the user to read the instructions and get ready to start. The average time to perform the experiment is 8 minutes.

4.1.4 Data acquisition

EEG is recorded using a set of 16 electrodes (F_z, FC_3, FC_z, FC_4, T_7, C_3, C_z, C_4, T_8, TP_7, CP_3, CP_z, CP_4, CP_8, PO_3, PO_4) placed according to International 10-20 system, referenced to the ear and with N_z as ground (Figure 4.4).

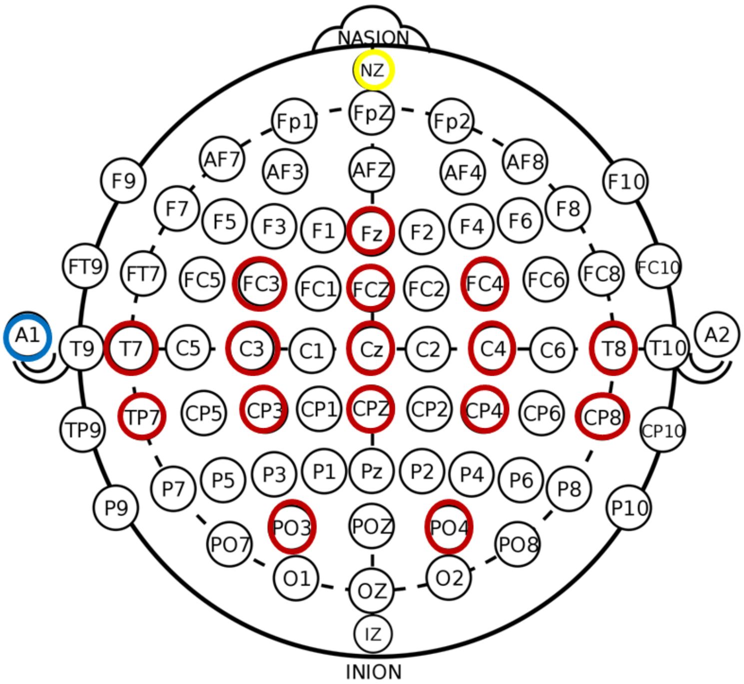


Figure 4.4: Electrodes displacement.

The stimulus presentation, the online BCI system and the offline analyses are implemented in Matlab2018b (MathWorks), making use of the Lab Streaming Layer (LSL) library for signal acquisition and Psychophysics Toolbox Version 3 for the interfaces.

4.1.5 Data processing

The process starts with playing the sequence of sounds, that is previously determined by randomly ordering the stimuli. While each sound is played, samples of the recorded signal are acquired, as well as their timestamps, and stored in a matrix, that is updated every time a new sound in played, by concatenations of samples. The signals are sampled at 256 Hz. A Notch filter with band 48 - 52 Hz and a bandpass 5th order Butterworth filter between 0.5 and 40 Hz are applied.

Each stimulus corresponds to a vector of 102 samples ($(sound + pause) \times samplingfrequency$) with the addition of 77 samples (300 ms) in order to consider the period that follows the end of the stimulus, that is when the P300 occurs (see Figure 4.5). All the data corresponding to the same stimulus are aligned and averaged, in order to obtain one mean vector of samples for each stimulus. Before doing that, the first repetition of each stimulus is discarded. During the first repetition the user is still not really familiar with the sounds, affecting the signal with this kind of confusion. A block diagram of the whole process is shown in Figure 4.6.

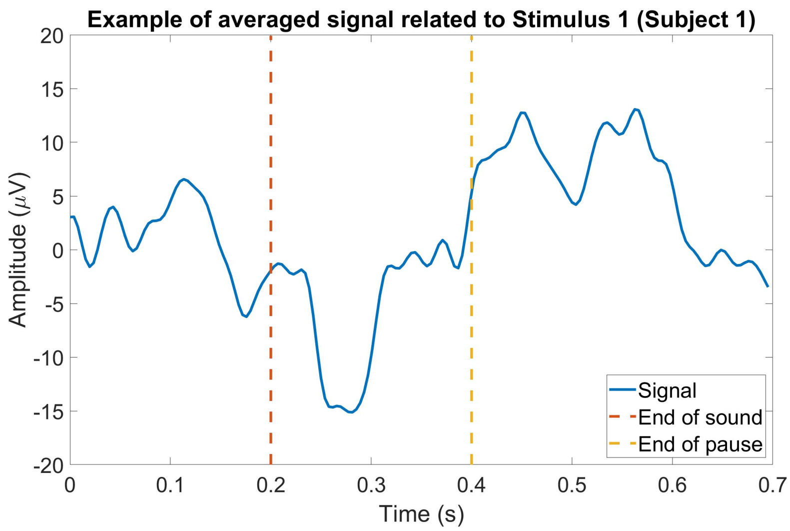


Figure 4.5: Example of averaged signal related to Stimulus 1 (Subject 1).

Features extraction is performed in order to reduce the dimensionality of the dataset and to enhance signal characteristics. Features have been extracted both from the signal itself and from its power spectral density (PSD), that is computed applying *Welch's PSD estimator*. Given a signal $x[n]$, assumed to be wide sense stationary (WSS), the PSD is defined as the

Fourier transform of its auto-correlation sequence:

$$P_{x,x}(f) = T \sum_{m=-\infty}^{+\infty} r_{x,x}[m] e^{-j2\pi fmT} \tag{4.1}$$

Matlab function *pwelch* returns PSD estimate (*pxx*) of an input signal and the related frequencies *f*, using Welch's overlapped segment averaging estimator:

$$[pxx, f] = pwelch(x, window, noverlap, nfft, fs) \tag{4.2}$$

where x is the input signal, *window* is the dimension of the segments to consider for the estimation, *noverlap* is the number of overlapped samples between adjacent segments, *nfft* is the number of points for the Discrete Fourier Transform (DFT) and *fs* is the sampling frequency. The following features have been extracted:

1. *Maximum*: maximum value of the signal.
2. *Minimum*: minimum value of the signal.
3. *Mean*: mean value of the signal.
4. *Skewness*: a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the centre point.
5. *Kurtosis*: a measure of whether the data are peaked or flat relative to a normal distribution.
6. *Mobility*: represents the mean frequency or the proportion of standard deviation of the power spectrum.
7. *Complexity*: gives an estimate of the bandwidth of the signal, which indicates the similarity of the shape of the signal to a pure sine wave.
8. *Shannon entropy*: the measure of information entropy associated with each possible data value.
9. *Sample entropy*: a measure of the complexity of physiological time-series signals.
10. *Higher peak in P300 range*: maximum value of the signal in the interval 200 - 500 ms after the end of the stimulus.

11. *Time of higher peak in P300 range*: time of occurrence of Feature 10.

12. *Maximum of PSD*: maximum value of *pxx*.

13. *Frequency of maximum of PSD*: frequency of occurrence of Feature 12.

14. *Energy*: estimation of the energy of the signal as the integral of *pxx*.

15. *Hurst exponent*: a measure of long-term memory of time series. It relates to the autocorrelations of the signal.

16. *Maximum of cross-correlation with a target signal*: measure of similarity between the signal and a target signal (Figure 4.5).

17. *Position of maximum of cross-correlation with a target signal*: position of 16.

Table 4.2 shows the features divided in categories.

Statistical	**Probabilistic**	**Time Domain**	**Frequency Domain**	**Geometrical**
Maximum Minimum Mean Skewness Kurtosis Mobility Complexity	Shannon entropy Sample entropy	Higher peak in P300 range Time of higher peak in P300 range	Maximum of PSD Frequency of maximum of PSD Energy	Hurst exponent Maximum of cross-correlation with a target signal Position of maximum of cross-correlation with a target signal

Table 4.2: Features extracted from the signal.

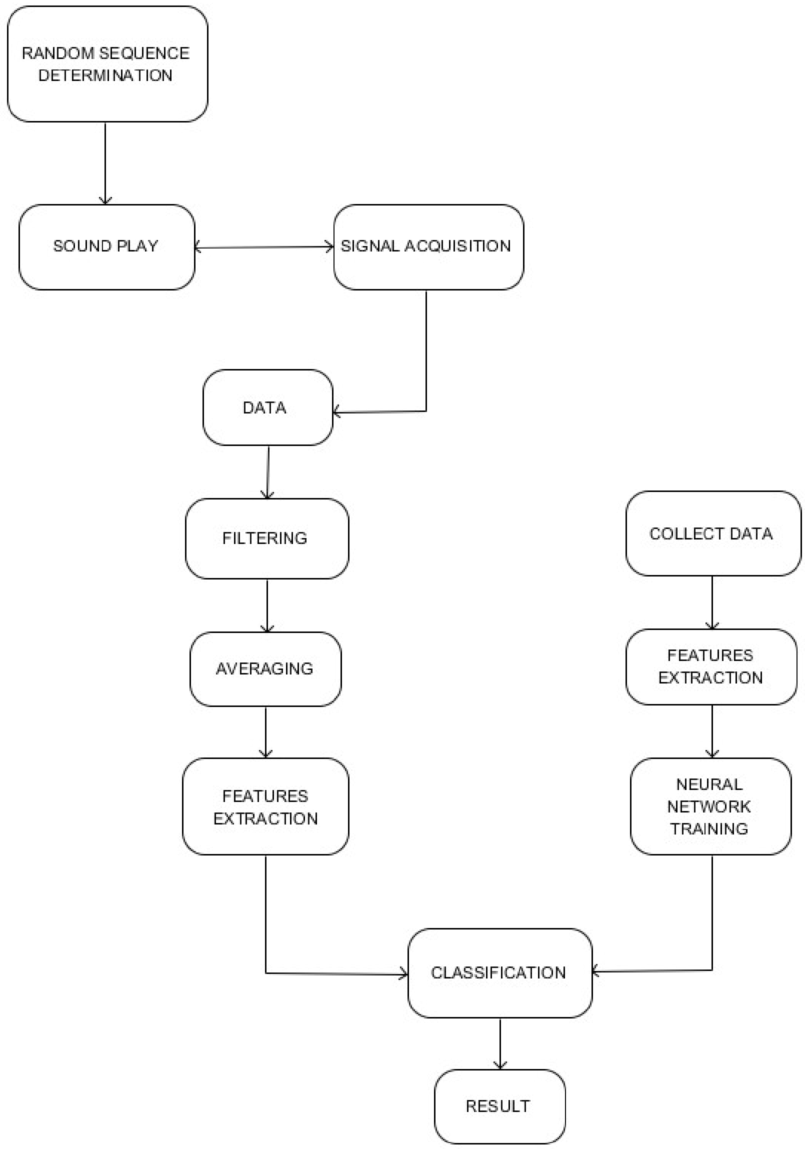


Figure 4.6: Block diagram of data processing.

4.1.6 Signal classification

The classification of the signal is performed by a NN, implemented in Matlab with the *Deep Learning Toolbox*. The NN is trained on features extracted from the signal and classifies them using the *Softmax function* as activation function. Softmax function normalizes its input into a probability distribution, such that every element x_i in the output is in the interval [0,1] and

$\sum_i x_i = 1$. The goal is to obtain a binary classification: 1 for the target stimulus, 0 for the non-target ones. The classification is unbalanced because, as mentioned before, in a sequence of stimuli the ratio between target and non-targets is 1:6. For taking into account classes imbalance, the NN must perform a weighted classification: 86% and 14% for target and non-target respectively.

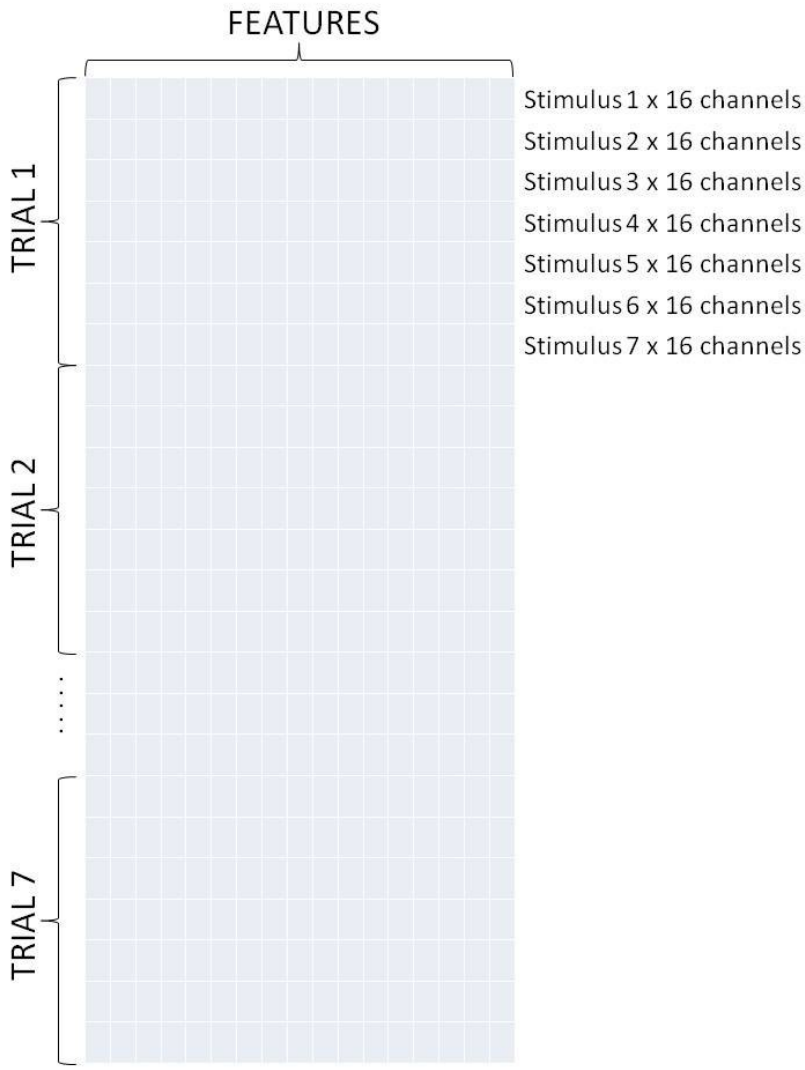


Figure 4.7: Structure of the training set.

The NN is composed by the following layers:

- *Input layer.*

- *Fully connected layers* with dimensions 15, 10 and 2.

- *Softmax layer.*

- *Weighted classification layer.*

The training of the NN has been performed on a dataset of signals from the participants to the experiment. The training set is a $M \times N$ matrix, where $M = 16\, channels \times 7\, stimuli \times 7\, trials \times number\ of\ sessions$, and $N = number\ of\ features$. Each row of the matrix represents the input for the Neural Network (Figure 4.7).

The classification carries out the following steps:

- *Threshold step*: class *target* is assigned if the probability of that class resulting from the softmax function is higher than the threshold, class *non-target* is selected otherwise.

- *Score step*: a score is assigned to each stimulus, by counting for how many channels the signal was in class *target*. Select as target the stimulus with the higher score.

- *Comparison step*: this step is applied in case of a tie (i.e. where there is more than one winner from the previous steps). In this case, class *target* is assigned if the probability of that class resulting from the softmax function is higher than the probability of class *non-target*, that is selected otherwise.

The flow chart of the classification process is shown in Figure 4.8.

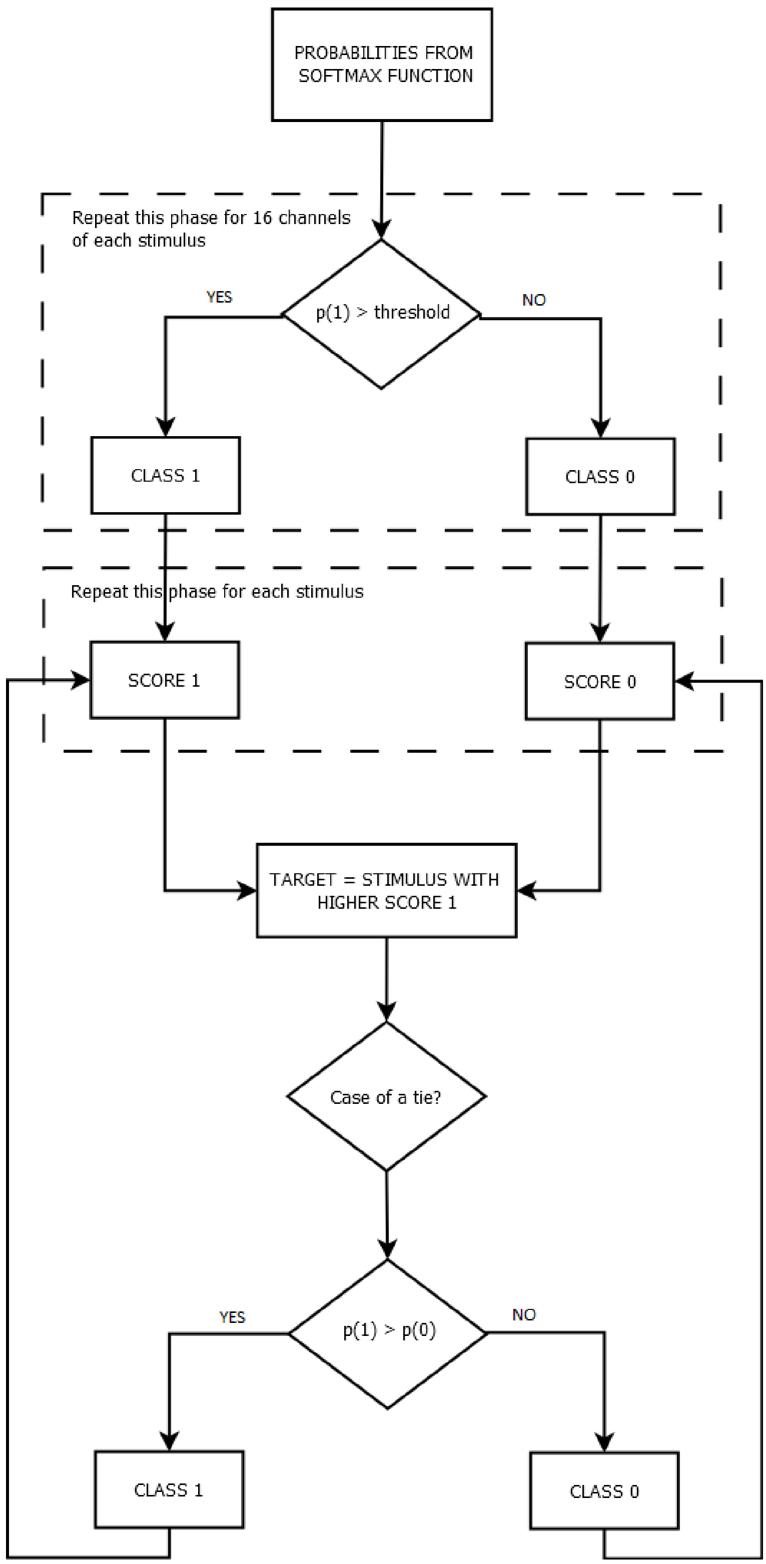


Figure 4.8: Flowchart of the classification process.

4.2 ET system

The ET system adds control features to the spelling system described above. It lets the user delete an unwanted letter or save a word and starting a new one.

4.2.1 Participants

Ten healthy subjects (5 males, 5 females, age 26.01±3.20 years) participated in the experiments. All subjects were volunteering students or researchers. No one had previous experiences with ET systems.

4.2.2 Equipment

The ET system includes an eye tracker connected to the computer. The main components of the eye tracker (Eye Tribe) are a camera and a high-resolution infrared LED (Table 4.3).

Sampling rate	30 Hz and 60 Hz mode
Accuracy	0.5° (average)
Spatial resolution	0.1° (RMS)
Latency	$<$ 20 ms at 60 Hz
Calibration	5, 9, 12 points
Operating range	45 – 75 cm
Tracking area	40 $\times$ 30 cm at 65 cm distance
Screen sizes	Up to 24 inches
API/SDK	C++, C# and Java
Data output	Binocular gaze data
Dimensions	(W/H/D) 20 $\times$ 1.9 $\times$ 1.9 cm (7.9 $\times$ 0.75 $\times$ 0.75 inches)
Weight	70 g
Connection	USB 3.0 Superspeed

Table 4.3: Data sheet of the Eye Tribe tracker. Adapted from [109].

The Eye Tribe's device uses the camera to track the user's eye movement. The camera tracks even the most minuscule of movements of the users' pupils, by taking the images and running them through computer-vision algorithms (Figure 4.9).

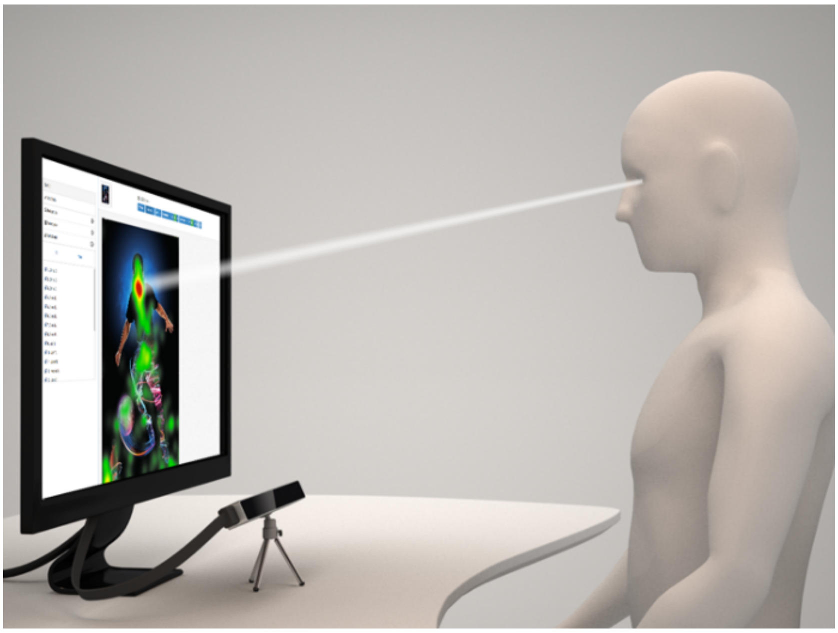


Figure 4.9: The Eye Tribe tracker. Reprinted from [109].

4.2.3 Experimental procedure

The ET system locates the user's gaze position on the screen. It represents the control part of the overall system, because it lets the user delete an unwanted letter or save the current word and start a new one. This is possible by detecting two positions of the user's gaze - *up* and *down* - and assigning all the other positions to *center*.

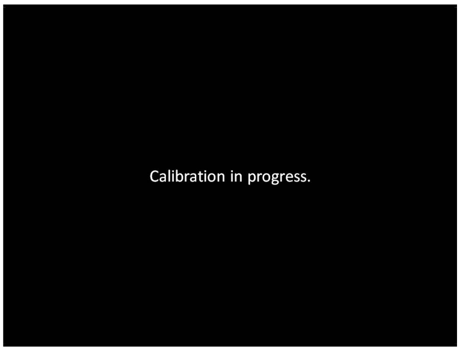
The procedure starts with the *calibration*. This step is useful to set the screen coordinates and display the regions referred to each position. During calibration process two white rectangles are shown on a black screen, first up then down, and the user has to look in that regions for two seconds each. The regions included by the rectangles are 250 pixel high and screen-like wide.

In Figure 4.10 there are examples of the interfaces shown during the calibration step.

(a) Interface 1: beginning of calibration.

(b) Interface 2: position *up*.

(c) Interface 3: position *down*.

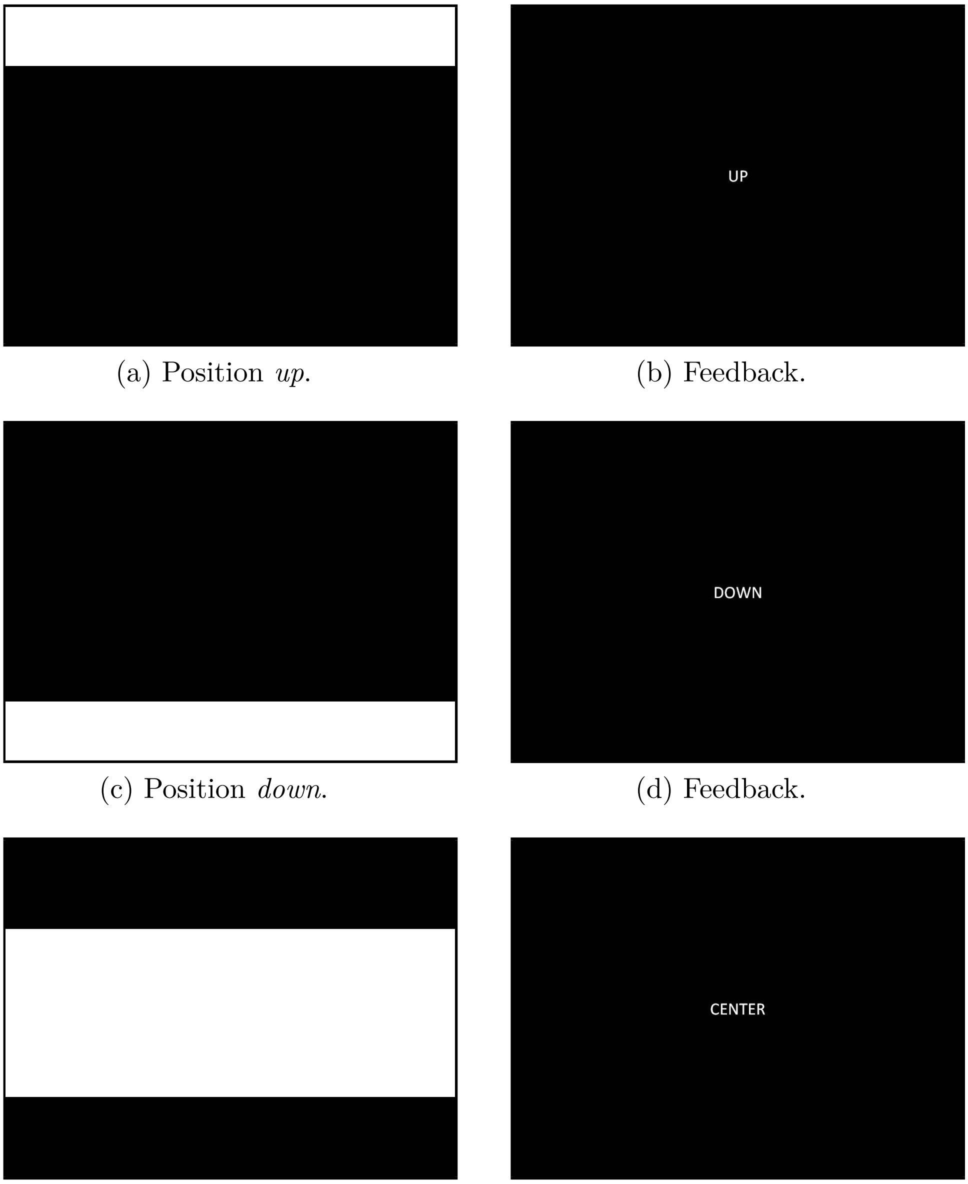


(d) Interface 4: calibration in progress.

Figure 4.10: Example of interfaces shown during the calibration step.

The procedure includes 30 trials: in each trial a rectangle is shown (up, down or center) in random order. The regions included by the *up* and *down* rectangles are 250 pixel high and screen-like wide, while the *center* rectangle is 1000 pixel high. The user has to look at the shown region for 1 second. Then the interface shows the detected position as feedback (Figure 4.11).

The rectangles must be considered as the lowest and the highest boundaries for the up and down regions respectively. This means that the up region includes all the space from the bottom side of the up rectangle upward, and the down region includes all the space from the top side of the down rectangle downward. For this reason, looking at the computer screen when selecting up or down is not mandatory.



(a) Position *up*. (b) Feedback.

(c) Position *down*. (d) Feedback.

(e) Position *center*. (f) Feedback.

Figure 4.11: Example of interfaces shown during the experimental procedure.

4.2.4 Region classification

Region classification - *up*, *down* or *center* - is performed with a threshold method. The calibration process returns the maximum and minimum y coordinate of the gaze - max_y and min_y respectively - so that it is possible to establish a threshold for regions discrimination. The classification process takes into account the last y coordinate of the gaze in each trial. It returns the following:

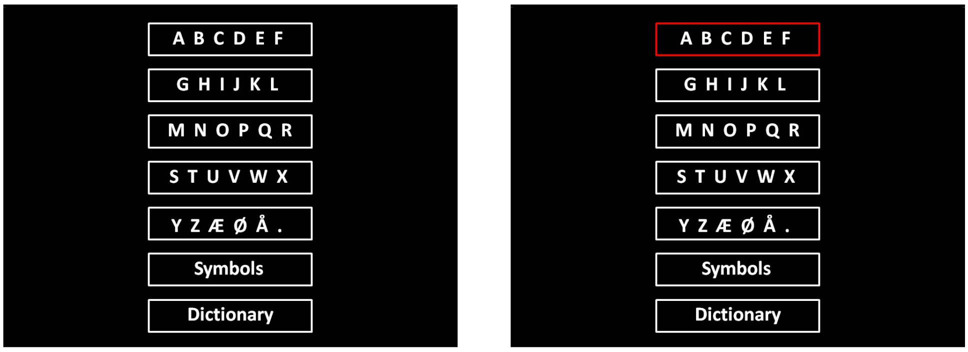
- UP: $y < (min_y + 250)\, pixel$.
- DOWN: $y > (max_y - 250)\, pixel$.
- CENTER: $(max_y - 250\, pixel) < y < (min_y + 250)\, pixel$.

The x coordinate is not taken into account.

4.3 Hybrid system

The hybrid spelling system employs the BCI to let the user select a letter from a GUI and the ET as control system to delete and unwanted letter and save a word. The system is completely hand-free and the user is asked to face the computer screen where a GUI is presented.

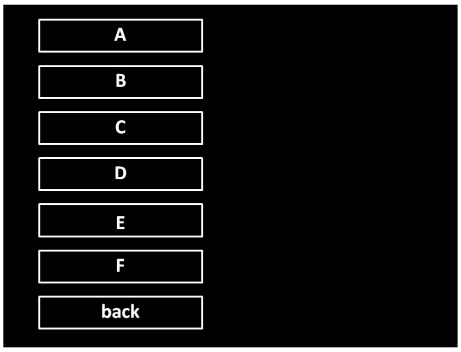
The first interface shows 7 boxes containing letters and symbols (Figure 4.12a). Each box is related to a sound, in the same order described in 4.1.3. At the beginning, the boxes are highlighted in red sequentially while the related sound is played: in this way the user knows and remembers which sound he has to concentrate on in order to select that box (Figure 4.12b). This happens every time a new interface is presented.



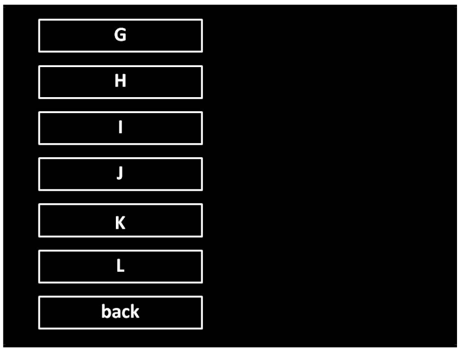
(a) Interface 1: Main window. (b) Interface 2: Main window when associating boxes and sounds.

Figure 4.12: Initial interface of the spelling system.

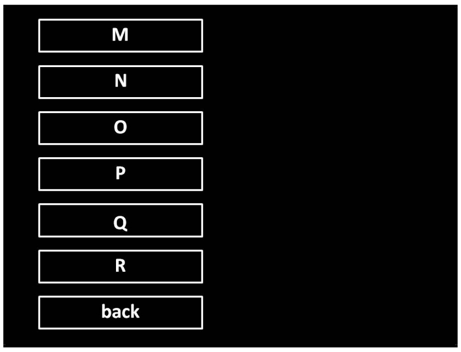
The procedure is as set out in Section 4.1.3: a sequence of 7 sounds repeated 20 times in random order is played. The user has to concentrate on the sound associated to the box he wants to select in order to access it. The following interface is related to the chosen box as shown in Figure 4.13.



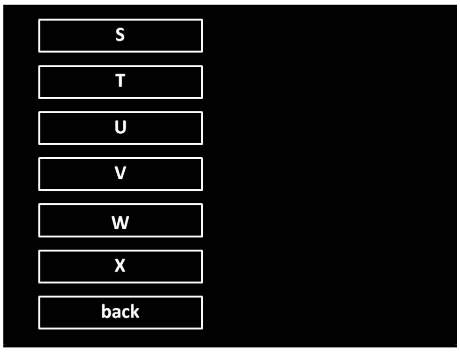
(a) Interface 3: Window related to the first box.



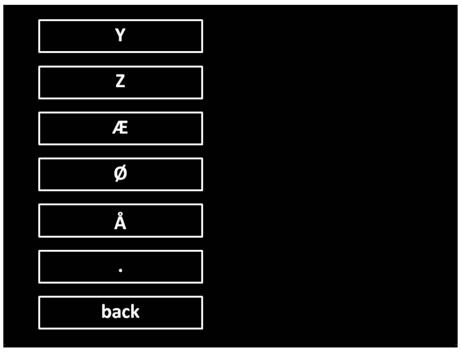
(b) Interface 4: Window related to the second box.



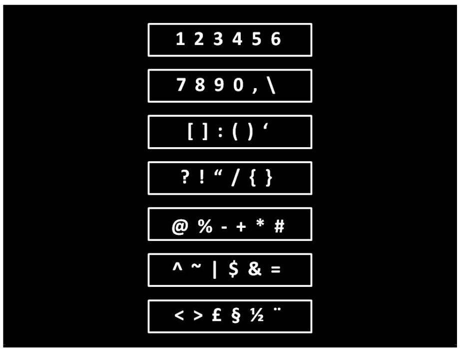
(c) Interface 5: Window related to the third box.



(d) Interface 6: Window related to the fourth box.



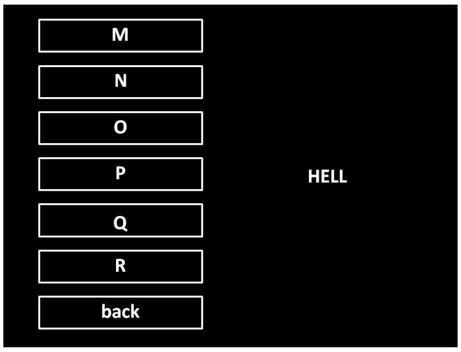
(e) Interface 7: Window related to the fifth box.



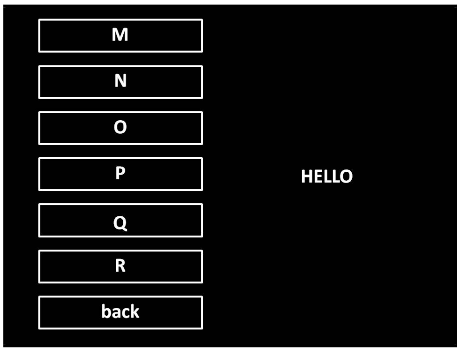
(f) Interface 8: Window related to the sixth box.

Figure 4.13: Interfaces presented after choosing a box from the initial interface.

When Interfaces 3 to 7 (Figure 4.13a - to 4.13e) are presented, the procedure is as above. If a current word is present, it appears on the right side of the interface (Figure 4.14a). When selecting a new letter (boxes 1 to 6) it appears next to the current word (Figure 4.14b), before coming back to the initial interface (Figure 4.12a). If 'back' is selected (box 7), the system comes back to the initial interface without adding letters to the current word.



(a) Interface 9: Window related to letter selection with current word.

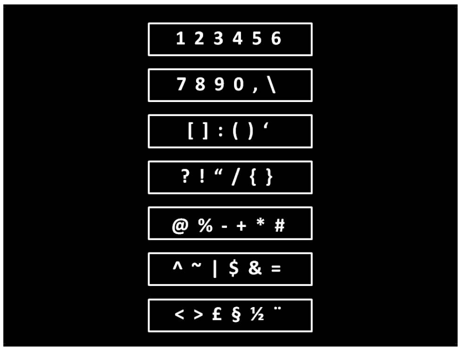


(b) Interface 10: Window related to letter addition to current word.

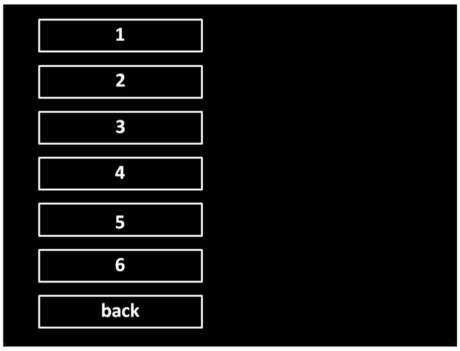
Figure 4.14: Letter selection and addition to current word.

If 'Symbols' (box 6) from the initial interface is selected, there is an extra step before selecting the target symbol (Figure 4.15).

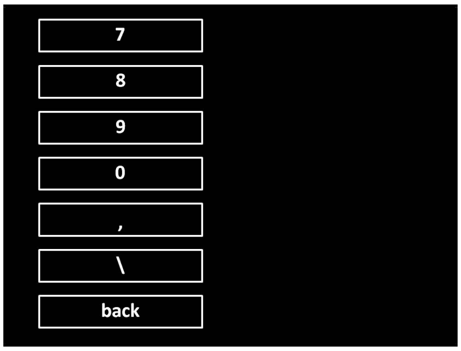
After selecting a letter or a symbol, the system always comes back to the initial interface (Figure 4.12a). Before repeating the same procedure to select another character, there is a pause of 1 second. During this period the ET takes action: if the user looks up the last letter is deleted, if he looks down the current word is saved on a .txt file and a new word stars, if he looks on the central region of the screen the procedure goes on as before. The ET works only during this period and only when the main window is presented. It is not possible to delete a letter or to save a word when another interface is presented or when the sounds are played.



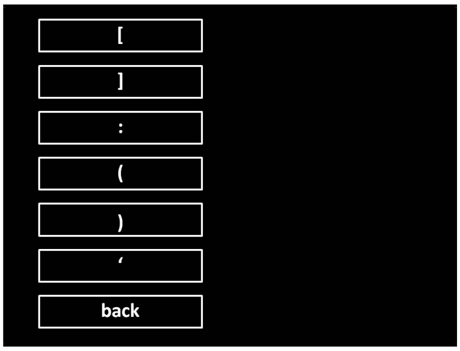
(a) Interface 8: Window related to 'Symbol' box.



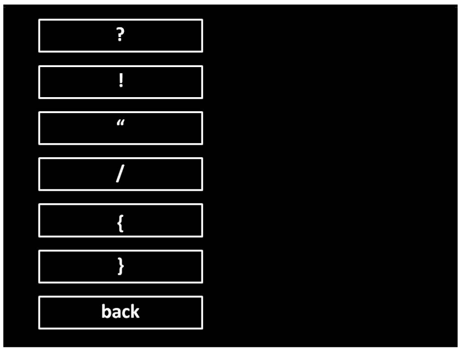
(b) Interface 11: Window related to the first box of Interface 8.



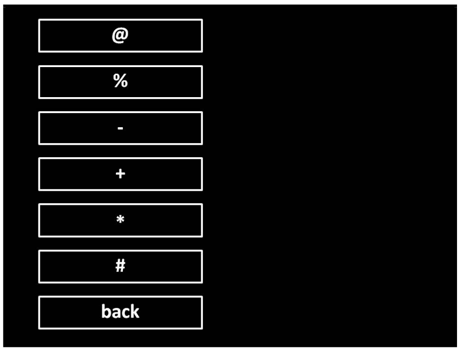
(c) Interface 12: Window related to the second box of Interface 8.



(d) Interface 13: Window related to the third box of Interface 8.

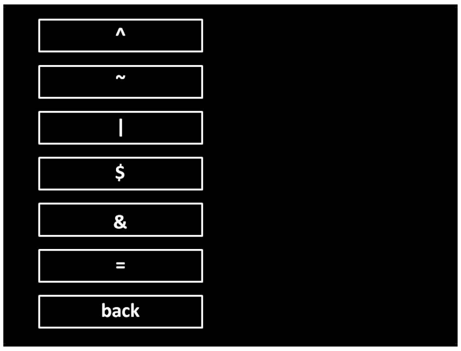


(e) Interface 14: Window related to the fourth box of Interface 8.

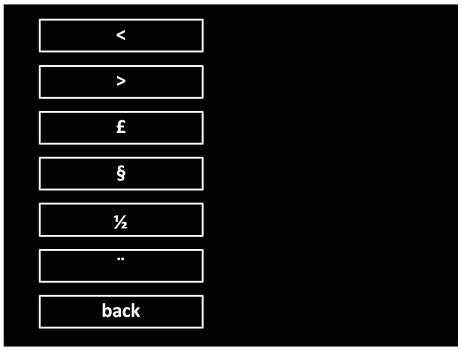


(f) Interface 15: Window related to the fifth box of Interface 8.

Figure 4.15: Interfaces presented after choosing the 'Symbols' box from the initial interface.



(g) Interface 16: Window related to the sixth box of Interface 8.



(h) Interface 17: Window related to the seventh box of Interface 8.

Figure 4.15: Interfaces presented after choosing the 'Symbols' box from the initial interface.

Chapter 5

Results

This chapter reports experimental results used to verify the performances of the methods described in the previous chapter.

5.1 BCI system

The performances of the BCI system are related to the results of the classification process. It includes the Neural Network and a further step that takes into account the channels of the signal acquisition.

5.1.1 Neural Network results

The classification has been performed using a training set that includes 210 trials from all the subjects of the test set. The performances of the NN can be evaluated from the confusion matrix - shown in Figure 5.1 - where class 1 correspond to the target stimulus and class 0 to the non-target ones.

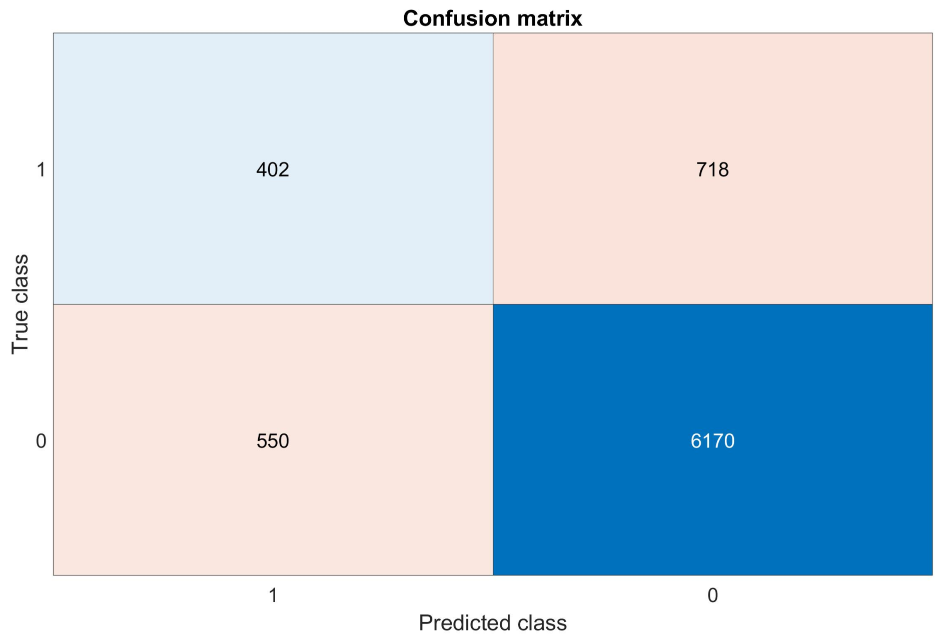


Figure 5.1: Confusion matrix of the NN classification.

$$Sensitivity\ (True\ Positive\ Rate) = \frac{TP}{TP + FN} = 36\% \tag{5.1}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} = 84\% \tag{5.2}$$

$$Specificity\ (True\ Negative\ Rate) = \frac{TN}{TN + FP} = 92\% \tag{5.3}$$

Even if the accuracy is high, the sensitivity is quite low with respect to the specificity, meaning that the NN classifies better the non-target stimuli with respect to the target. This is due to the unbalance of the classes.

5.1.2 Classification results

The most important index for the classification performances of this experiment is the *sensitivity*, since the final goal is to identify the target stimulus, represented by the *True class*.

The NN step is followed by other steps in order to improve the sensitivity, as described in Section 4.1.6. The output of the NN is represented by a probability of belonging to the classes. It is possible to define a threshold for selecting the right class, considering the probability distribution of the two classes (Figure 5.2).

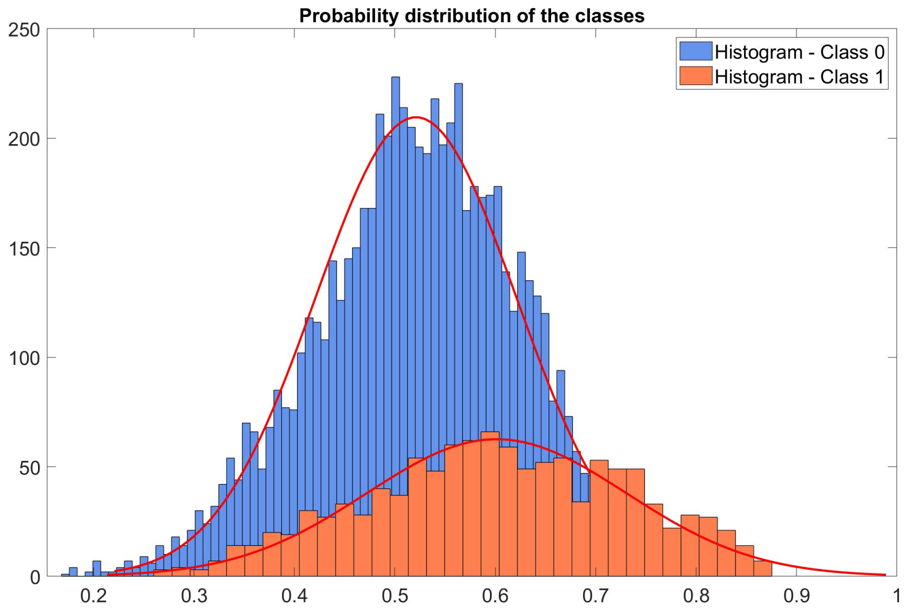


Figure 5.2: Probability distribution of the classes.

Even if the probability curves are not completely separated, Class 1 prevails over Class 0 after the intersection point (~0.7). This represents the best choice for the threshold.

Once the threshold has been established, the classification process assigns Class 1 to the signals whose probability is higher than the threshold itself and chooses as target the signal that has more ones over the 16 channels. This is useful to increase the sensitivity as it introduces a further step that takes the channels into account. The results of the whole classification process are shown in Figures 5.3 and 5.4.

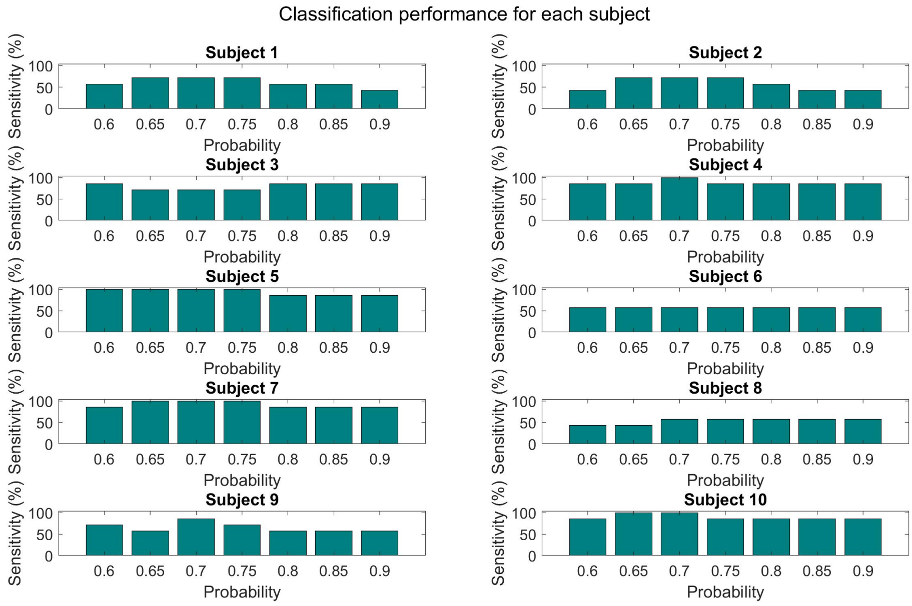


Figure 5.3: Sensitivity of each subject with respect to different probability thresholds.

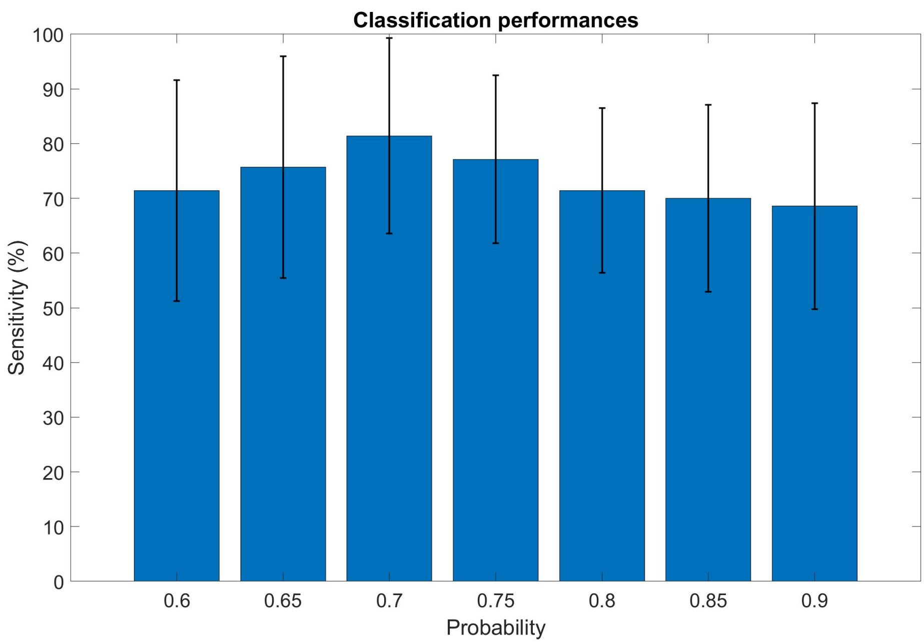


Figure 5.4: Mean sensitivity with respect to different probability thresholds.

Considering 0.7 as the best threshold, the classification reaches a final sensitivity of 81%.

5.2 ET system

The performances of the ET system have been evaluated both for single region classification and for the overall classification. The single-region performances are useful to understand which is the best classified region, while the average performances are related to the whole system.

Accuracy is used as a statistical measure of how well the classification test correctly identifies the target region. It is the proportion of true results (both true positives and true negatives) among the total number of cases examined.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.4}$$

5.2.1 Single-region accuracy

The experiment includes three regions that the user can select: *up*, *down* or *center*. The regions are presented 30 times - 10 times each - in random order. The accuracy of the classification of each region is calculated as the true results of the region among the total number of trials when that region is presented (10).

$$Accuracy_{\mathrm{UP}} = \frac{(TP + TN)_{up}}{10} \tag{5.5}$$

$$Accuracy_{\mathrm{DOWN}} = \frac{(TP + TN)_{down}}{10} \tag{5.6}$$

$$Accuracy_{\mathrm{CENTER}} = \frac{(TP + TN)_{center}}{10} \tag{5.7}$$

The results are shown in Table 5.1.

Subject	Accuracy (UP)	Accuracy (DOWN)	Accuracy (CENTER)
1	100%	90%	100%
2	100%	90%	100%
3	100%	100%	100%
4	100%	100%	100%
5	100%	100%	100%
6	100%	100%	100%
7	100%	100%	100%
8	90%	100%	100%
9	100%	100%	100%
10	100%	100%	100%
Average	99%	98%	100%

Table 5.1: Single-region accuracy for each subject.

All the regions have been well classified at least 9 times out of 10 for every subject. The best classified region is *center* (100%), while *up* and *down* have a similar accuracy (99% and 98% respectively).

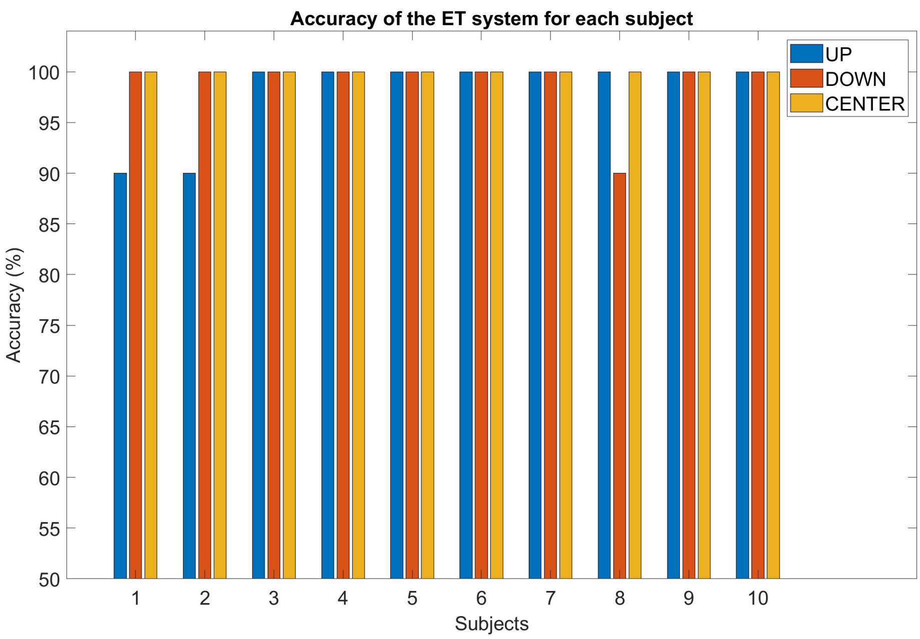


Figure 5.5: Bar diagrams of single-region accuracy for each subject.

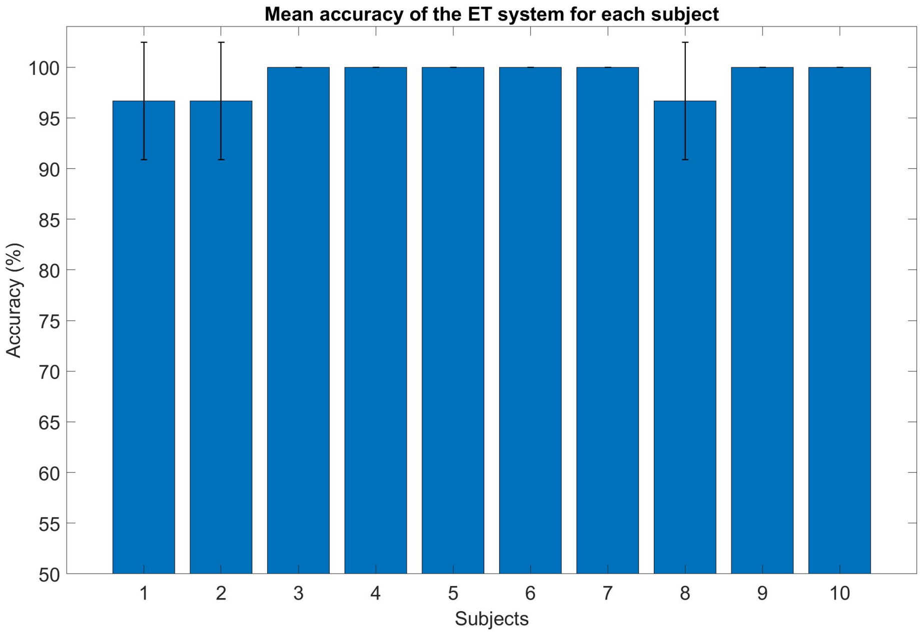


Figure 5.6: Bar diagrams of mean accuracy for each subject.

5.2.2 Mean accuracy

The mean accuracy represents the performance of the overall experiment. It is calculated for each subject:

$$Accuracy_n = \frac{(TP + TN)_n}{30} \tag{5.8}$$

Table 5.2 shows that the lowest accuracy is 97% (Subjects 1 - 2 - 8) and the highest is 100% (all the other Subjects).

Subject	Accuracy
1	97%
2	97%
3	100%
4	100%
5	100%
6	100%
7	100%
8	97%
9	100%
10	100%
Average	99%

Table 5.2: Mean accuracy for each subject.

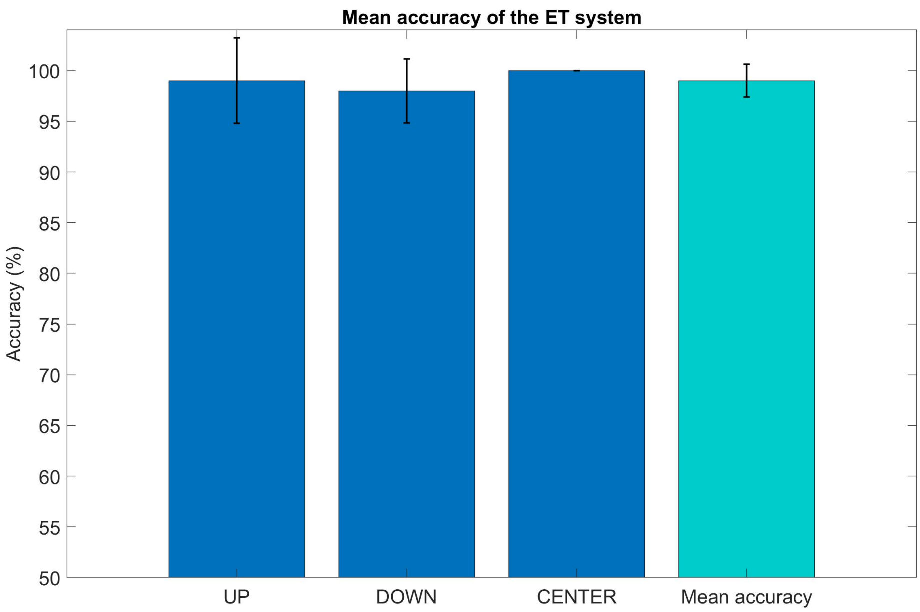


Figure 5.7: Bar diagrams of mean accuracy.

The mean accuracy of the whole experiment is calculated as:

$$Accuracy = \frac{\sum_{n=1}^{10} \frac{(TP+TN)_n}{30}}{10} \tag{5.9}$$

The ET system has an accuracy of 99%.

Chapter 6

Discussion

This chapter starts from commenting the results obtained from the current model, and continues with describing the possible methods that have been applied to the model in order to try to improve the results. Furthermore, some ideas for future works on this project are presented.

6.1 Comments on the current results

The analysis of the performances of the current system shows good results. Considering the *sensitivity* as the most important indicator for the BCI experiment, it goes from 36% after the first classification step to 81% at the end of the whole classification process. This result can be considered good enough but it can also be improved.

An important issue for this experiments concerns the time needed to select each character. In order to evaluate the time, it necessary to consider a sequence of 7 sounds, lasting 400 ms (200 ms for the sound and 200 ms for the pause), each repeated 20 times, for a single selection. This means that the whole process should be repeated twice in order to select a letter, since the first selection is related to the group of letters (see Section 4.3).

$$Time = (0.2 + 0.2) \times 7 \times 20 \times 2 = 112\, seconds \tag{6.1}$$

Equation 6.1 shows that the time needed for each character selection is 112 seconds. This results must be improved for a real use of the system.

6.2 Possible improvements

In this section, three methods for trying to improve the current results are presented. New results will be showed after applying these methods to the initial model.

6.2.1 Balanced dataset

Classes imbalance in the current dataset may have influence on the classification performances - mostly on *sensitivity* - that represents the most relevant index for the BCI experiment. Dealing with this condition is possible through an "adaptive synthetic (ADASYN) sampling approach". The essential idea of ADASYN is to use a weighted distribution for different minority class examples according to their level of difficulty in learning, where more synthetic data is generated for minority class examples that are harder to learn compared to those minority examples that are easier to learn [110].

The goal is to increase the initial sensitivity of the NN in order to improve the final performance. After applying the ADASYN method, the dataset becomes balanced and nearly doubled, as shown in Table 6.1.

	Before ADASYN	**After ADASYN**
Class 0	86%	50%
Class 1	14%	50%
Dataset	23520	40328

Table 6.1: Dataset before and after the application of ADASYN.

Analyzing the performances of the NN classification, it is possible to evaluate that the accuracy is the same as before. About the sensitivity - that is, as repeatedly mentioned, a key parameter in this context - it is slightly higher with respect to the initial model (Figure 6.1 and Table 6.2).

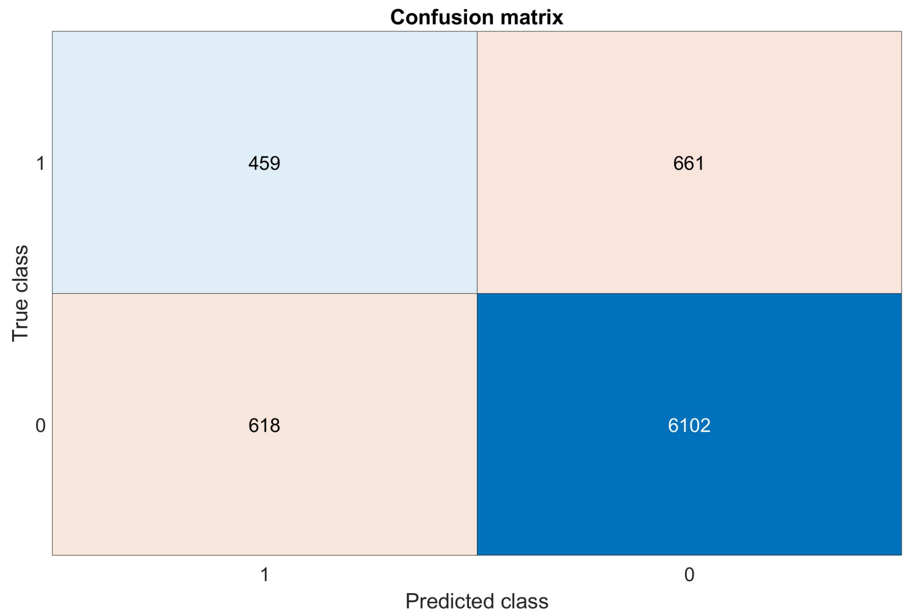


Figure 6.1: Confusion matrix of the NN classification after the application of ADASYN.

	Before ADASYN	**After ADASYN**
Sensitivity	36%	43%
Accuracy	84%	84%
Specificity	92%	90%

Table 6.2: Classification performances of the NN before and after the application of ADASYN.

According to the results shown in Figures 6.2 and 6.3, the overall sensitivity of the classification is 78%, approximately equal to the sensitivity of the initial model. The lowest sensitivity is 55% (Subject 6) and the highest is 100% (Subjects 5 and 7).

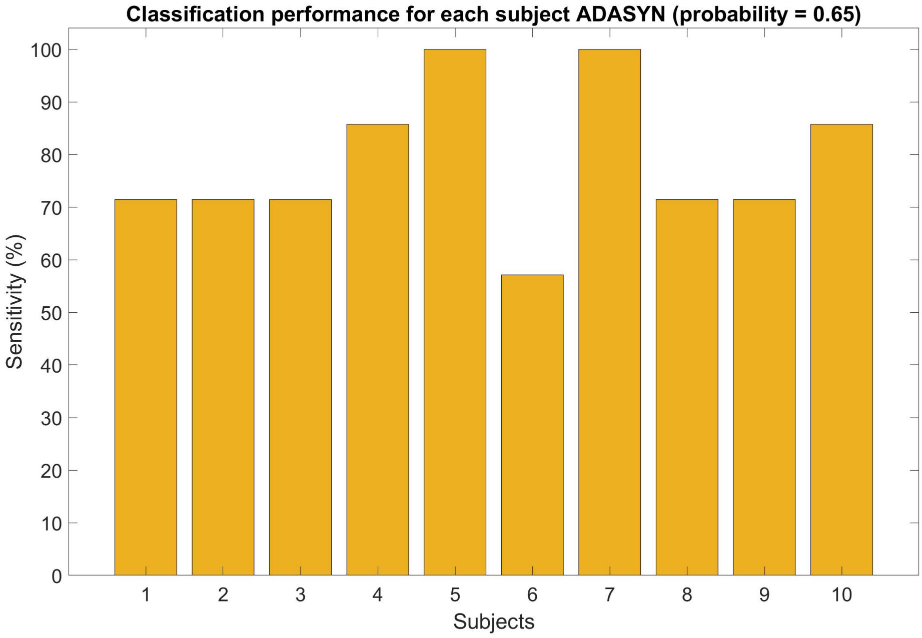


Figure 6.2: Sensitivity of each subject after the application of ADASYN.

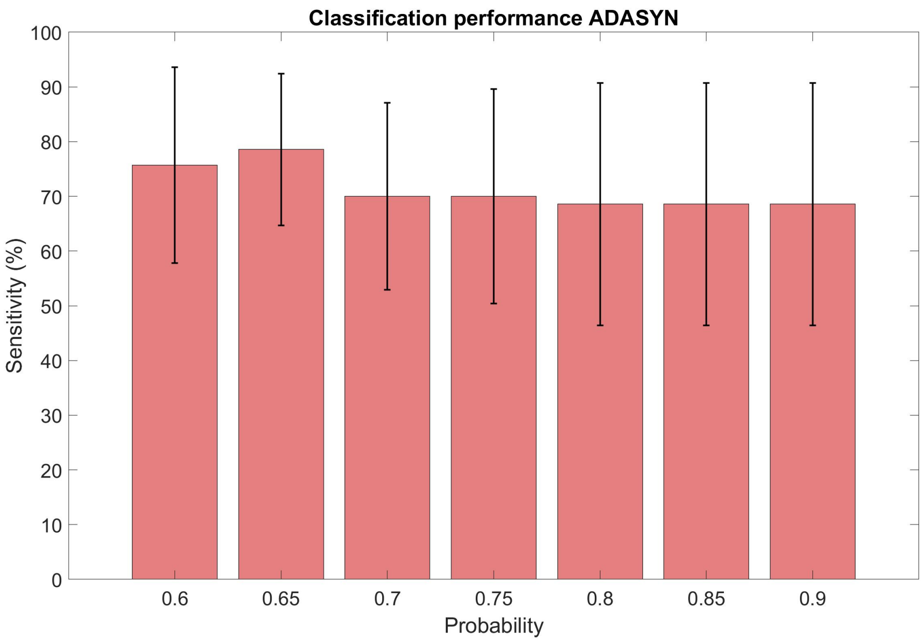


Figure 6.3: Mean sensitivity with respect to different probability thresholds after the application of ADASYN.

Comparing the performances of the classification before and after the application of the ADASYN method, it is not possible to state that a balanced dataset with this method leads to tangible improvements in the model. The performances are comparable.

6.2.2 Feature selection

Feature selection aims at creating an accurate predictive model by choosing features that give as good or better accuracy whilst requiring less data. It is useful to identify and remove unneeded, irrelevant and redundant attributes from data that do not contribute to the accuracy of the model or may in fact decrease the accuracy of the model. It also aids to reduce the complexity of the model and the time for the classification.

Feature selection has been performed by starting from the first feature and sequentially adding the others one by one. In this way 17 models have been obtained - each including from 1 to 17 features respectively.

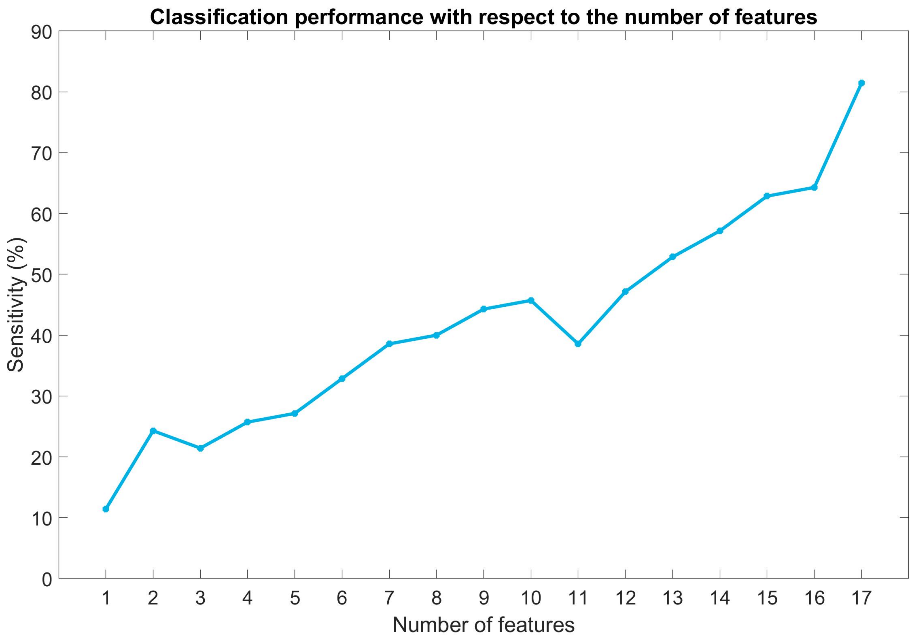


Figure 6.4: Sensitivity of the model with respect to the number of features.

According to Figure 6.4, the highest sensitivity is obtained with 17 features, which represent the initial model. Using less features would reduce the sensitivity of the classification. Extracting more features will probability increase the performance, but such method has not been implemented in the current project.

6.2.3 Time reduction

Time reduction represents an important issue for this system. According to Equation 6.1, each character selection takes 112 seconds. This result depends on the following factors:

- number of stimuli
- number of repetitions
- length of each stimulus (sounds + pause).

Each factor needs to be reduced in order to reduce the selection time. Using less stimuli will mean to simplify the system by including only letters in the interface. A lower number of repetitions would be possible for a trained user, that is able to concentrate on the target stimulus also if he can hear it less times. The length of each stimulus can also be considered a matter of training or ability of the user to distinguish different sounds even if they are played for a short time.

Experiments have been performed on 3 subjects, using a model that includes 5 stimuli, each one lasting 150 ms with 150 ms of pause, and 10 repetitions. In this way the time for selecting one character is reduced from 112 seconds to:

$$Time = (0.15\ +\ 0.15) \times 5 \times 10 \times 2 = 30\,seconds \tag{6.2}$$

This would represents a significant improvement for the system, since the needed time is less one third of the initial time.

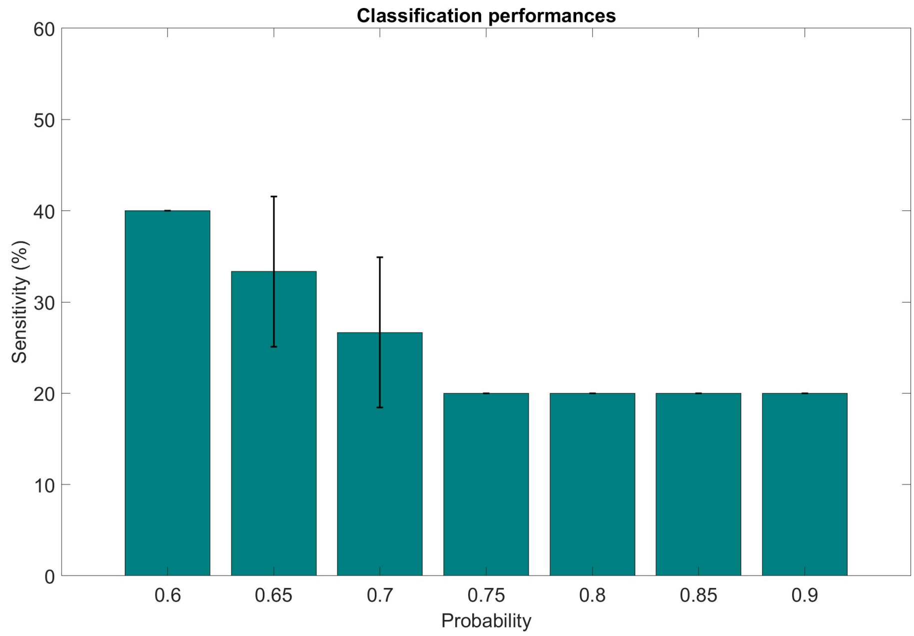


Figure 6.5: Classification performances of the system with time reduction.

Figure 6.5 shows the sensitivity of the classification process applied to the new model. The highest value (40%) is achieved by all the subjects with a probability threshold of 0.6. The result is not successful as it was for the initial system (81%), but it has been obtained using the same NN employed for the initial model. This network is trained on data related to the initial model, meaning that the result cannot be really accurate. The classification performances may be improved by using a NN that is trained with data related to this model.

6.3 Future work

This section will provide an overview of some ideas for improving the performances of the model in a future work.

6.3.1 Dataset

The dataset plays a key role in the classification performances. The current training set includes 210 trials from all the subjects that took part in the experiments. It is possible to understand the influence of the dataset comparing the classification performances obtained using different training sets.

Since it is not possible to obtain a bigger dataset in the current experiment, the comparison has been performed using smaller datasets. The properties of each dataset are shown in Table 6.3.

Training set	Number of subjects	Number of trials
1	1	56
2	5	105
3	10	210

Table 6.3: Properties of the training sets used for the comparison.

In the first model, the training set includes 56 trials from only one subject. It includes 105 trials from half of the subjects in the second model. The third model is the one used for the whole experiment.

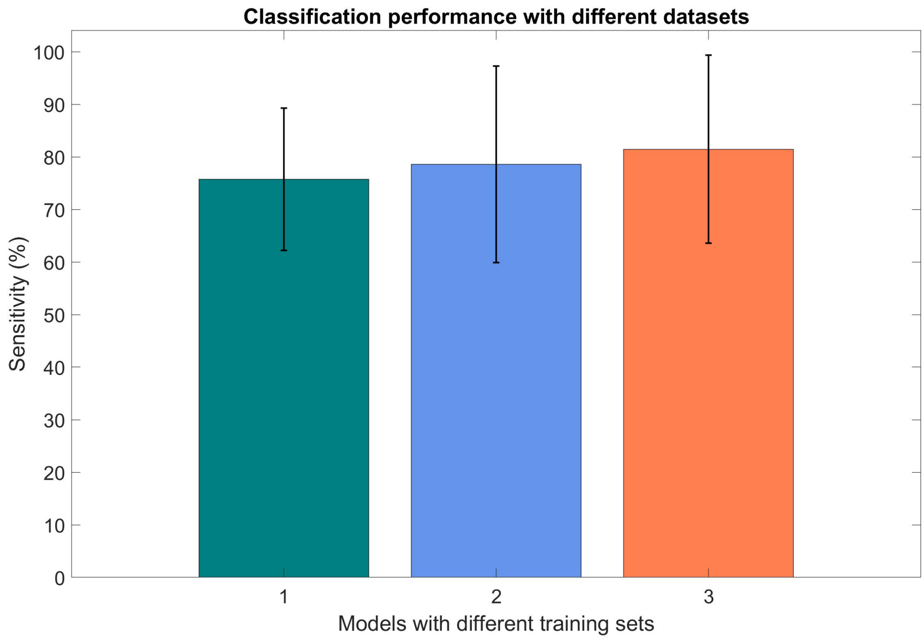


Figure 6.6: Comparison of the classification sensitivity obtained using 3 different datasets.

The sensitivity of the classification (Figure 6.6) increases as the size of the dataset increases. It is 76% for model 1, 79% for model 2 and 81% for model 3. Even though it is not a big difference, it may result in a significant improvement if the dataset was much bigger or if it included many more subjects at least for the training of the NN.

6.3.2 Parameters tuning

The BCI experiment is build by combining a signal acquisition and processing phase with a classification phase. Both processes are parameter-dependent, since they involve the choice of many features that can be tuned, maybe resulting in different results.

The signal acquisition phase includes parameters as *sound duration*, *pause duration*, *number of repetitions*, *type of sounds*, *electrodes displacement.* These parameters have been tuned on one subject, trying to find a combination that leads to better results. Many combinations of these parameters may be tested on different subjects, or on the same subject if he/she would be the actual user of the system.

The multilayer perceptron NN involves parameters as *number of layers*, *number of neurons*, *number of epochs*, *bach size*, *optimizer*, *activation function.* The choice of Softmax as activation function is reported in Section 4.1.6. All the other parameters have been tuned in order to find a combination between good performances and low time consumption. Also in this case, many methods can be applied for tuning the parameters, to find, perhaps, a better combination.

6.3.3 System improvements

The system employed in this experiments includes a laptop computer that plays the sounds, and 16 electrodes that record the EEG signal.

The direction of the sounds plays an important role in the stimuli discrimination from the user. It adds a new feature - the *orientation* - that helps the user to concentrate on the target sound since he knows which direction the stimulus comes from. Many experiments have been conducted, placing speakers in different direction, with the same distance from the user. In [111] [112] [113] [114] they used respectively 8, 5, 3 and 10 sound directions, one for each stimulus, obtaining better results with respect to the single-direction experiments.

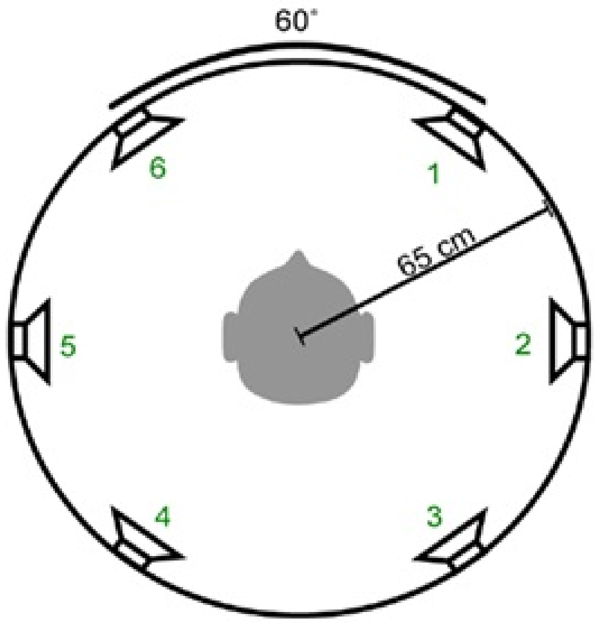


Figure 6.7: Sound orientation in the BCI experiment. Adapted from [115].

Another limit of the equipment employed in this project is the number of electrodes. As the results presented in Section 5.1.2 show, the performance of the classification are highly improved when considering the results obtained by each of the 16 electrodes. For this reason, using more electrodes - as in [116] [117] (more than 32) and [115] [118] (more than 56) - may represent a gain for the system performances.

6.3.4 System application

The final goal of this system is to be employed by people affected by LIS as an assitive technology. Future works may improve the performances in order to build a spelling system that can actually help these people.

Chapter 7

Conclusion

The goal of this project was to build an assistive technology that combines two methods for human-machine interaction: a Brain Computer Interface and an Eye Tracking system. The final result is a completely hand-free hybrid spelling system for people affected by severe disabilities.

Good performances have been reached, represented by a 81% sensitivity for the BCI and a 99% accuracy for the ET system.

Many methods have been applied to try to improve the classification performances for the BCI - that represents the main part of the system - but none of them was able to increase the sensitivity, which is a crucial parameter in this case.

The most critical issue was the time needed for the selection of each character - 112 seconds - that is not sustainable for a real application of this system. A future work is certainly needed to improve this parameter, as well as to improve the classification performances by using a bigger dataset for the training of the Neural Network.

Bibliography

[1] R. Z. M. Chan, *World report on disability*, `http://apps.who.int/iris/bitstream/handle/10665/70670/WHO_NMH_VIP_11.01_eng.pdf;jsessionid=E84543FD4CBC5080865D7CED40298D53?sequence=1`, [Online; accessed 19-September-2018], 2013.

[2] B. S. Armour, E. A. Courtney-Long, M. H. Fox, H. Fredine, and A. Cahill, "Prevalence and causes of paralysis—united states, 2013", *American journal of public health*, vol. 106, no. 10, pp. 1855–1857, 2016.

[3] M. Obermann and M. Lyon, "Financial cost of amyotrophic lateral sclerosis: A case study", *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, vol. 16, no. 1-2, pp. 54–57, 2015.

[4] R. G. Lupu, F. Ungureanu, and V. Siriteanu, "Eye tracking mouse for human computer interaction", in *E-Health and Bioengineering Conference (EHB), 2013*, IEEE, 2013, pp. 1–4.

[5] R. Spataro, M. Ciriacono, C. Manno, and V. La Bella, "The eye-tracking computer device for communication in amyotrophic lateral sclerosis", *Acta Neurologica Scandinavica*, vol. 130, no. 1, pp. 40–45, 2014.

[6] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review", *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012.

[7] H. Drewes, "Eye gaze tracking for human computer interaction", PhD thesis, lmu, 2010.

[8] E. Margalit, J. Weiland, R. Clatterbuck, G. Fujii, M. Maia, M. Tameesh, G. Torres, S. D'Anna, S. Desai, D. Piyathaisere, *et al.*, "Visual and electrical evoked response recorded from subdural electrodes implanted above the visual cortex in normal dogs under two methods of anesthesia", *Journal of neuroscience methods*, vol. 123, no. 2, pp. 129–137, 2003.

[9] Z. C. Chao, Y. Nagasaka, and N. Fujii, “Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkey”, *Frontiers in neuroengineering*, vol. 3, p. 3, 2010.

[10] G. Schalk, “Can electrocorticography (ecog) support robust and powerful brain-computer interfaces?”, *Frontiers in neuroengineering*, vol. 3, p. 9, 2010.

[11] E. Ryapolova-Webb, P. Afshar, S. Stanslaski, T. Denison, C. de Hemptinne, K. Bankiewicz, and P. A. Starr, “Chronic cortical and electromyographic recordings from a fully implantable device: Preclinical experience in a nonhuman primate”, *Journal of neural engineering*, vol. 11, no. 1, p. 016 009, 2014.

[12] C. S. Mestais, G. Charvet, F. Sauter-Starace, M. Foerster, D. Ratel, and A. L. Benabid, “Wimagine: Wireless 64-channel ecog recording implant for long term clinical applications”, *IEEE transactions on neural systems and rehabilitation engineering*, vol. 23, no. 1, pp. 10–21, 2015.

[13] T. Matsuo, K. Kawasaki, T. Osada, H. Sawahata, T. Suzuki, M. Shibata, N. Miyakawa, K. Nakahara, A. Iijima, N. Sato, *et al.*, “Intrasulcal electrocorticography in macaque monkeys with minimally invasive neurosurgical protocols”, *Frontiers in systems neuroscience*, vol. 5, p. 34, 2011.

[14] K. A. Davis, B. K. Sturges, C. H. Vite, V. Ruedebusch, G. Worrell, A. B. Gardner, K. Leyde, W. D. Sheffield, and B. Litt, “A novel implanted device to wirelessly record and analyze continuous intracranial canine eeg”, *Epilepsy research*, vol. 96, no. 1-2, pp. 116–122, 2011.

[15] J. J. Howbert, E. E. Patterson, S. M. Stead, B. Brinkmann, V. Vasoli, D. Crepeau, C. H. Vite, B. Sturges, V. Ruedebusch, J. Mavoori, *et al.*, “Forecasting seizures in dogs with naturally occurring epilepsy”, *PloS one*, vol. 9, no. 1, e81920, 2014.

[16] G. Schalk, J. Kubanek, K. Miller, N. Anderson, E. Leuthardt, J. Ojemann, D. Limbrick, D. Moran, L. Gerhardt, and J. Wolpaw, “Decoding two-dimensional movement trajectories using electrocorticographic signals in humans”, *Journal of neural engineering*, vol. 4, no. 3, p. 264, 2007.

[17] S. Engel, T. Aksenova, and A. Eliseyev, “Kernel-based npls for continuous trajectory decoding from ecog data for bci applications”, in *International Conference on Latent Variable Analysis and Signal Separation*, Springer, 2017, pp. 417–426.

[18] X. Pei, D. L. Barbour, E. C. Leuthardt, and G. Schalk, “Decoding vowels and consonants in spoken and imagined words using electrocorticographic signals in humans”, *Journal of neural engineering*, vol. 8, no. 4, p. 046 028, 2011.

[19] C. Potes, A. Gunduz, P. Brunner, and G. Schalk, “Dynamics of electrocorticographic (ecog) activity in human temporal and frontal cortical areas during music listening”, *Neuroimage*, vol. 61, no. 4, pp. 841–848, 2012.

[20] C. Potes, P. Brunner, A. Gunduz, R. T. Knight, and G. Schalk, “Spatial and temporal relationships of electrocorticographic alpha and gamma activity during auditory processing”, *Neuroimage*, vol. 97, pp. 188–195, 2014.

[21] P. Brunner, K. Dijkstra, W. G. Coon, J. Mellinger, A. L. Ritaccio, and G. Schalk, “An ecog-based bci based on auditory attention to natural speech”, in *Brain-Computer Interface Research*, Springer, 2017, pp. 7–19.

[22] X. Pei, E. C. Leuthardt, C. M. Gaona, P. Brunner, J. R. Wolpaw, and G. Schalk, “Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition”, *Neuroimage*, vol. 54, no. 4, pp. 2960–2972, 2011.

[23] J. Kubanek, P. Brunner, A. Gunduz, D. Poeppel, and G. Schalk, “The tracking of speech envelope in the human cortex”, *PloS one*, vol. 8, no. 1, e53398, 2013.

[24] W. Wang, J. L. Collinger, A. D. Degenhart, E. C. Tyler-Kabara, A. B. Schwartz, D. W. Moran, D. J. Weber, B. Wodlinger, R. K. Vinjamuri, R. C. Ashmore, *et al.*, “An electrocorticographic brain interface in an individual with tetraplegia”, *PloS one*, vol. 8, no. 2, e55344, 2013.

[25] G. Schalk and E. C. Leuthardt, “Brain-computer interfaces using electrocorticographic signals”, *IEEE reviews in biomedical engineering*, vol. 4, pp. 140–154, 2011.

[26] S. Ryun, J. S. Kim, D. Lee, and C. K. Chung, “Prediction of motor and somatosensory function from human ecog”, in *Brain-Computer Interface (BCI), 2018 6th International Conference on*, IEEE, 2018, pp. 1–4.

[27] H. Choi, J. Lee, J. Park, B. H. Cho, K.-M. Lee, and D. P. Jang, "Movement state classification for bimanual bci from non-human primate's epidural ecog using three-dimensional convolutional neural network", in *Brain-Computer Interface (BCI), 2018 6th International Conference on*, IEEE, 2018, pp. 1–3.

[28] T. Yanagisawa, M. Hirata, Y. Saitoh, H. Kishima, K. Matsushita, T. Goto, R. Fukuma, H. Yokoi, Y. Kamitani, and T. Yoshimine, "Electrocorticographic control of a prosthetic arm in paralyzed patients", *Annals of neurology*, vol. 71, no. 3, pp. 353–361, 2012.

[29] E. S. Nurse, S. E. John, D. R. Freestone, T. J. Oxley, H. Ung, S. F. Berkovic, T. J. O'Brien, M. J. Cook, and D. B. Grayden, "Consistency of long-term subdural electrocorticography in humans", *IEEE Trans Biomed Eng*, 2017.

[30] G. Buzsáki, C. A. Anastassiou, and C. Koch, "The origin of extracellular fields and currents—eeg, ecog, lfp and spikes", *Nature reviews neuroscience*, vol. 13, no. 6, p. 407, 2012.

[31] V. S. Polikov, P. A. Tresco, and W. M. Reichert, "Response of brain tissue to chronically implanted neural electrodes", *Journal of neuroscience methods*, vol. 148, no. 1, pp. 1–18, 2005.

[32] A. P. Georgopoulos, A. B. Schwartz, and R. E. Kettner, "Neuronal population coding of movement direction", *Science*, vol. 233, no. 4771, pp. 1416–1419, 1986.

[33] A. B. Schwartz, "Motor cortical activity during drawing movements: Population representation during sinusoid tracing", *Journal of neurophysiology*, vol. 70, no. 1, pp. 28–36, 1993.

[34] J. Wessberg, C. R. Stambaugh, J. D. Kralik, P. D. Beck, M. Laubach, J. K. Chapin, J. Kim, S. J. Biggs, M. A. Srinivasan, and M. A. Nicolelis, "Real-time prediction of hand trajectory by ensembles of cortical neurons in primates", *Nature*, vol. 408, no. 6810, p. 361, 2000.

[35] D. M. Taylor, S. I. H. Tillery, and A. B. Schwartz, "Direct cortical control of 3d neuroprosthetic devices", *Science*, vol. 296, no. 5574, pp. 1829–1832, 2002.

[36] M. Velliste, S. Perel, M. C. Spalding, A. S. Whitford, and A. B. Schwartz, "Cortical control of a prosthetic arm for self-feeding", *Nature*, vol. 453, no. 7198, p. 1098, 2008.

[37] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt, *et al.*, “Reach and grasp by people with tetraplegia using a neurally controlled robotic arm”, *Nature*, vol. 485, no. 7398, p. 372, 2012.

[38] B. Jarosiewicz, N. Y. Masse, D. Bacher, S. S. Cash, E. Eskandar, G. Friehs, J. P. Donoghue, and L. R. Hochberg, “Advantages of closed-loop calibration in intracortical brain–computer interfaces for people with tetraplegia”, *Journal of neural engineering*, vol. 10, no. 4, p. 046 012, 2013.

[39] F. R. Willett, C. Pandarinath, B. Jarosiewicz, B. A. Murphy, W. D. Memberg, C. H. Blabe, J. Saab, B. L. Walter, J. A. Sweet, J. P. Miller, J. M. Henderson, K. V. Shenoy, J. D. Simeral, L. R. Hochberg, R. F. Kirsch, and A. B. Ajiboye, “Feedback control policies employed by people using intracortical brain–computer interfaces”, *Journal of Neural Engineering*, vol. 14, no. 1, p. 016 001, 2017. [Online]. Available: `http://stacks.iop.org/1741-2552/14/i=1/a=016001`.

[40] T. Milekovic, “Brain-computer interfaces based on intracortical recordings of neural activity for restoration of movement and communication of people with paralysis”, in *Brain-Computer Interface (BCI), 2018 6th International Conference on*, IEEE, 2018, pp. 1–2.

[41] B. Gunasekera, T. Saxena, R. Bellamkonda, and L. Karumbaiah, “Intracortical recording interfaces: Current challenges to chronic recording function”, *ACS Chemical Neuroscience*, vol. 6, no. 1, pp. 68–83, 2015, PMID: 25587704. DOI: `10.1021/cn5002864`. [Online]. Available: `https://doi.org/10.1021/cn5002864`.

[42] B. D. Winslow, M. B. Christensen, W.-K. Yang, F. Solzbacher, and P. A. Tresco, “A comparison of the tissue response to chronically implanted parylene-c-coated and uncoated planar silicon microelectrode arrays in rat cortex”, *Biomaterials*, vol. 31, no. 35, pp. 9163–9172, 2010.

[43] A. J. Woolley, H. A. Desai, and K. J. Otto, “Chronic intracortical microelectrode arrays induce non-uniform, depth-related tissue responses”, *Journal of neural engineering*, vol. 10, no. 2, p. 026 007, 2013.

[44] D. I. Kim, J. Sui, S. Rachakonda, T. White, D. S. Manoach, V. P. Clark, B.-C. Ho, S. C. Schulz, and V. D. Calhoun, “Identification

of imaging biomarkers in schizophrenia: A coefficient-constrained independent component analysis of the mind multi-site schizophrenia study", *Neuroinformatics*, vol. 8, no. 4, pp. 213–229, 2010.

[45] T. L. Richards and V. W. Berninger, "Abnormal fmri connectivity in children with dyslexia during a phoneme task: Before but not after treatment", *Journal of neurolinguistics*, vol. 21, no. 4, pp. 294–304, 2008.

[46] R. G. Wise and C. Preston, "What is the value of human fmri in cns drug development?", *Drug discovery today*, vol. 15, no. 21-22, pp. 973–980, 2010.

[47] S. Ogawa and T.-M. Lee, "Magnetic resonance imaging of blood vessels at high fields: In vivo and in vitro measurements and image simulation", *Magnetic resonance in medicine*, vol. 16, no. 1, pp. 9–18, 1990.

[48] S. Ogawa, T.-M. Lee, A. S. Nayak, and P. Glynn, "Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields", *Magnetic resonance in medicine*, vol. 14, no. 1, pp. 68–78, 1990.

[49] S. Ogawa, T.-M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation", *Proceedings of the National Academy of Sciences*, vol. 87, no. 24, pp. 9868–9872, 1990.

[50] S. Ogawa, R. Menon, D. Tank, S. Kim, H. Merkle, J. Ellermann, and K. Ugurbil, "Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. a comparison of signal characteristics with a biophysical model", *Biophysical journal*, vol. 64, no. 3, pp. 803–812, 1993.

[51] N. Weiskopf, R. Veit, M. Erb, K. Mathiak, W. Grodd, R. Goebel, and N. Birbaumer, "Physiological self-regulation of regional brain activity using real-time functional magnetic resonance imaging (fmri): Methodology and exemplary data", *Neuroimage*, vol. 19, no. 3, pp. 577–586, 2003.

[52] A. Eklund, M. Andersson, H. Ohlsson, A. Ynnerman, and H. Knutsson, "A brain computer interface for communication using real-time fmri", in *Pattern Recognition (ICPR), 2010 20th International Conference on*, IEEE, 2010, pp. 3665–3669.

[53] S. Ruiz, K. Buyukturkoglu, M. Rana, N. Birbaumer, and R. Sitaram, "Real-time fmri brain computer interfaces: Self-regulation of single brain regions to networks", *Biological psychology*, vol. 95, pp. 4–20, 2014.

[54] M. O. Sokunbi, D. E. Linden, I. Habes, S. Johnston, and N. Ihssen, "Real-time fmri brain-computer interface: Development of a "motivational feedback" subsystem for the regulation of visual cue reactivity", *Frontiers in behavioral neuroscience*, vol. 8, p. 392, 2014.

[55] R. Sitaram, A. Caria, R. Veit, T. Gaber, S. Ruiz, and N. Birbaumer, "Volitional control of the anterior insula in criminal psychopaths using real-time fmri neurofeedback: A pilot study", *Frontiers in behavioral neuroscience*, vol. 8, p. 344, 2014.

[56] K. Buyukturkoglu, H. Roettgers, J. Sommer, M. Rana, L. Dietzsch, E. B. Arikan, R. Veit, R. Malekshahi, T. Kircher, N. Birbaumer, *et al.*, "Self-regulation of anterior insula with real-time fmri and its behavioral effects in obsessive-compulsive disorder: A feasibility study", *PloS one*, vol. 10, no. 8, e0135872, 2015.

[57] M. Stangl, G. Bauernfeind, J. Kurzmann, R. Scherer, and C. Neuper, "A haemodynamic brain–computer interface based on real-time classification of near infrared spectroscopy signals during motor imagery and mental arithmetic", *Journal of near infrared spectroscopy*, vol. 21, no. 3, pp. 157–171, 2013.

[58] B. Koo, H. Vu, H. Lee, H. Shin, and S. Choi, "Motor imagery detection with wavelet analysis for nirs-based bci", in *2016 4th International Winter Conference on Brain-Computer Interface (BCI)*, Feb. 2016, pp. 1–4. DOI: 10.1109/IWW-BCI.2016.7457441.

[59] L. C. Schudlo, S. D. Power, and T. Chau, "Dynamic topographical pattern classification of multichannel prefrontal nirs signals", *Journal of neural engineering*, vol. 10, no. 4, p. 046 018, 2013.

[60] H.-J. Hwang, J.-H. Lim, D.-W. Kim, and C.-H. Im, "Evaluation of various mental task combinations for near-infrared spectroscopy-based brain-computer interfaces", *Journal of Biomedical Optics*, vol. 19, no. 7, p. 077 005, 2014.

[61] C. Herff, D. Heger, O. Fortmann, J. Hennrich, F. Putze, and T. Schultz, "Mental workload during n-back task—quantified in the prefrontal cortex using fnirs", *Frontiers in human neuroscience*, vol. 7, p. 935, 2014.

[62] L. C. Schudlo and T. Chau, "Towards a ternary nirs-bci: Single-trial classification of verbal fluency task, stroop task and unconstrained rest", *Journal of neural engineering*, vol. 12, no. 6, p. 066 008, 2015.

[63] S. Waldert, T. Pistohl, C. Braun, T. Ball, A. Aertsen, and C. Mehring, "A review on directional information in neural signals for brain-machine interfaces", *Journal of Physiology-Paris*, vol. 103, no. 3-5, pp. 244–254, 2009.

[64] T. N. Lal, M. Schröder, N. J. Hill, H. Preissl, T. Hinterberger, J. Mellinger, M. Bogdan, W. Rosenstiel, T. Hofmann, N. Birbaumer, *et al.*, "A brain computer interface with online feedback based on magnetoencephalography", in *Proceedings of the 22nd international conference on Machine learning*, ACM, 2005, pp. 465–472.

[65] J. Zhang, G. Sudre, X. Li, W. Wang, D. J. Weber, and A. Bagic, "Clustering linear discriminant analysis for meg-based brain computer interfaces", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 3, pp. 221–231, 2011.

[66] N. I. Sabra and M. A. Wahed, "The use of meg-based brain computer interface for classification of wrist movements in four different directions", in *Radio Science Conference (NRSC), 2011 28th National*, IEEE, 2011, pp. 1–7.

[67] C. Reichert, S. Dürschmid, H.-J. Heinze, and H. Hinrichs, "A comparative study on the detection of covert attention in event-related eeg and meg signals to control a bci", *Frontiers in neuroscience*, vol. 11, p. 575, 2017.

[68] S. Baillet, "Magnetoencephalography for brain electrophysiology and imaging", *Nature neuroscience*, vol. 20, no. 3, p. 327, 2017.

[69] J. J. Vidal, "Toward direct brain-computer communication", *Annual Review of Biophysics and Bioengineering*, vol. 2, no. 1, pp. 157–180, 1973, PMID: 4583653. DOI: `10.1146/annurev.bb.02.060173.001105`. [Online]. Available: `https://doi.org/10.1146/annurev.bb.02.060173.001105`.

[70] B. Bromm, "Human brain electrophysiology. evoked potentials and evoked magnetic fields in science and medicine", *Pain*, vol. 39, no. 3, pp. 371–372, Dec. 1989. DOI: `10.1016/0304-3959(89)90058-4`. [Online]. Available: `https://doi.org/10.1016%2F0304-3959%2889%2990058-4`.

[71] M. Stadnyk, M. Fryz, and L. Scherbak, "The feature extraction and estimation of a steady-state visual evoked potential by the karhunen-loeve expansion", *Eastern-European Journal of Enterprise Technologies*, vol. 1, no. 4 (85), pp. 56–62, Feb. 2017. DOI: `10.15587/1729-4061.2017.91861`. [Online]. Available: `https://doi.org/10.15587%2F1729-4061.2017.91861`.

[72] A. Kolossa, T. Fingscheidt, K. Wessel, and B. Kopp, "A model-based approach to trial-by-trial p300 amplitude fluctuations", *Frontiers in Human Neuroscience*, vol. 6, 2013. DOI: `10.3389/fnhum.2012.00359`. [Online]. Available: `https://doi.org/10.3389%2Ffnhum.2012.00359`.

[73] J. Covington, "P300, stimulus intensity, and modality", *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, vol. 100, no. 6, pp. 579–584, Nov. 1996. DOI: `10.1016/s0921-884x(96)96013-0`. [Online]. Available: `https://doi.org/10.1016%2Fs0921-884x%2896%2996013-0`.

[74] C. C. Duncan, R. J. Barry, J. F. Connolly, C. Fischer, P. T. Michie, R. Näätänen, J. Polich, I. Reinvang, and C. V. Petten, "Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, p300, and n400", *Clinical Neurophysiology*, vol. 120, no. 11, pp. 1883–1908, Nov. 2009. DOI: `10.1016/j.clinph.2009.07.045`. [Online]. Available: `https://doi.org/10.1016%2Fj.clinph.2009.07.045`.

[75] M. Kerstin, W. Sarah, and S. Ute, "Neurofeedback of slow cortical potentials as an effective treatment of adult ADHD", *Frontiers in Human Neuroscience*, vol. 10, 2016. DOI: `10.3389/conf.fnhum.2016.220.00072`. [Online]. Available: `https://doi.org/10.3389%2Fconf.fnhum.2016.220.00072`.

[76] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor, "A spelling device for the paralysed", *Nature*, vol. 398, no. 6725, pp. 297–298, Mar. 1999. DOI: `10.1038/18581`. [Online]. Available: `https://doi.org/10.1038%2F18581`.

[77] N. Birbaumer, A. Kubler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, and H. Flor, "The thought translation device (TTD) for completely paralyzed patients", *IEEE Transactions on Rehabilitation Engineering*, vol. 8,

no. 2, pp. 190–193, Jun. 2000. DOI: 10.1109/86.847812. [Online]. Available: https://doi.org/10.1109%2F86.847812.

[78] N. Neumann, A. Kübler, J. Kaiser, T. Hinterberger, and N. Birbaumer, "Conscious perception of brain states: Mental strategies for brain–computer communication", *Neuropsychologia*, vol. 41, no. 8, pp. 1028–1036, Jan. 2003. DOI: 10.1016/s0028-3932(02)00298-1. [Online]. Available: https://doi.org/10.1016%2Fs0028-3932%2802%2900298-1.

[79] N. Birbaumer and F. Rodden, "The thought-translation device", in *Consciousness and Cognition*, Elsevier, 2007, pp. 69–IV. DOI: 10.1016/b978-012373734-2/50007-8. [Online]. Available: https://doi.org/10.1016%2Fb978-012373734-2%2F50007-8.

[80] I. Lazarou, S. Nikolopoulos, P. C. Petrantonakis, I. Kompatsiaris, and M. Tsolaki, "EEG-based brain–computer interfaces for communication and rehabilitation of people with motor impairment: A novel approach of the 21st century", *Frontiers in Human Neuroscience*, vol. 12, Jan. 2018. DOI: 10.3389/fnhum.2018.00014. [Online]. Available: https://doi.org/10.3389%2Ffnhum.2018.00014.

[81] R. Gibson, A. Owen, and D. Cruse, "Brain–computer interfaces for patients with disorders of consciousness", in *Progress in Brain Research*, Elsevier, 2016, pp. 241–291. DOI: 10.1016/bs.pbr.2016.04.003. [Online]. Available: https://doi.org/10.1016%2Fbs.pbr.2016.04.003.

[82] G. Pfurtscheller and A. Aranibar, "Event-related cortical desynchronization detected by power measurements of scalp EEG", *Electroencephalography and Clinical Neurophysiology*, vol. 42, no. 6, pp. 817–826, Jun. 1977. DOI: 10.1016/0013-4694(77)90235-8. [Online]. Available: https://doi.org/10.1016%2F0013-4694%2877%2990235-8.

[83] G. Pfurtscheller, "Event-related synchronization (ERS): An electrophysiological correlate of cortical areas at rest", *Electroencephalography and Clinical Neurophysiology*, vol. 83, no. 1, pp. 62–69, Jul. 1992. DOI: 10.1016/0013-4694(92)90133-3. [Online]. Available: https://doi.org/10.1016%2F0013-4694%2892%2990133-3.

[84] B. Blankertz, G. Dornhege, M. Krauledat, M. Schröder, J. Williamson, R. Murray-Smith, and K.-R. Müller, "The berlin brain-computer interface presents the novel mental typewriter hex-o-spell.", 2006.

[85] T. D'albis, R. Blatt, R. Tedesco, L. Sbattella, and M. Matteucci, "A predictive speller controlled by a brain-computer interface based on motor imagery", *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 19, no. 3, p. 20, 2012.

[86] S. Perdikis, R. Leeb, and J. d R Millán, "Context-aware adaptive spelling in motor imagery BCI", *Journal of Neural Engineering*, vol. 13, no. 3, p. 036 018, May 2016. DOI: `10.1088/1741-2560/13/3/036018`. [Online]. Available: `https://doi.org/10.1088%2F1741-2560%2F13%2F3%2F036018`.

[87] L. Cao, B. Xia, O. Maysam, J. Li, H. Xie, and N. Birbaumer, "A synchronous motor imagery based neural physiological paradigm for brain computer interface speller", *Frontiers in human neuroscience*, vol. 11, p. 274, 2017.

[88] S. Kotov, L. Turbina, P. Bobrov, A. Frolov, O. Pavlova, M. Kurganskaya, and E. Biryukova, "Rehabilitation of stroke patients with a bioengineered "brain–computer interface with exoskeleton" system", *Neuroscience and Behavioral Physiology*, vol. 46, no. 5, pp. 518–522, 2016.

[89] E. Biryukova, O. Pavlova, M. Kurganskaya, P. Bobrov, L. Turbina, A. Frolov, V. Davydov, A. Silchenko, and O. Mokienko, "Recovery of the motor function of the arm with the aid of a hand exoskeleton controlled by a brain–computer interface in a patient with an extensive brain lesion", *Human Physiology*, vol. 42, no. 1, pp. 13–23, 2016.

[90] A. A. Frolov, O. Mokienko, R. Lyukmanov, E. Biryukova, S. Kotov, L. Turbina, G. Nadareyshvily, and Y. Bushkova, "Post-stroke rehabilitation training with a motor-imagery-based brain-computer interface (BCI)-controlled hand exoskeleton: A randomized controlled multicenter trial", *Frontiers in Neuroscience*, vol. 11, Jul. 2017. DOI: `10.3389/fnins.2017.00400`. [Online]. Available: `https://doi.org/10.3389%2Ffnins.2017.00400`.

[91] J. Cantillo-Negrete, R. I. Carino-Escobar, P. Carrillo-Mora, D. Elias-Vinas, and J. Gutierrez-Martinez, "Motor imagery-based brain-computer interface coupled to a robotic hand orthosis aimed for neurorehabilitation of stroke patients", *Journal of Healthcare Engineering*, vol. 2018, pp. 1–10, 2018. DOI: `10.1155/2018/1624637`. [Online]. Available: `https://doi.org/10.1155%2F2018%2F1624637`.

[92] H. Wang and A. Bezerianos, "Brain-controlled wheelchair controlled by sustained and brief motor imagery BCIs", *Electronics Letters*, vol. 53, no. 17, pp. 1178–1180, Aug. 2017. DOI: 10.1049/el.2017.1637. [Online]. Available: https://doi.org/10.1049%2Fel.2017.1637.

[93] J. Zhuang and G. Yin, "Motion control of a four-wheel-independent-drive electric vehicle by motor imagery EEG based BCI system", in *2017 36th Chinese Control Conference (CCC)*, IEEE, Jul. 2017. DOI: 10.23919/chicc.2017.8028220. [Online]. Available: https://doi.org/10.23919%2Fchicc.2017.8028220.

[94] J. Tang, Y. Liu, D. Hu, and Z. Zhou, "Towards BCI-actuated smart wheelchair system", *BIOMEDICAL Engineering OnLine*, vol. 17, no. 1, Aug. 2018. DOI: 10.1186/s12938-018-0545-x. [Online]. Available: https://doi.org/10.1186%2Fs12938-018-0545-x.

[95] H. Chennamma and X. Yuan, "A survey on eye-gaze tracking techniques", *ArXiv preprint arXiv:1312.6410*, 2013.

[96] R. Soundariya and R. Renuga, "Eye movement based emotion recognition using electrooculography", in *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, IEEE, Apr. 2017. DOI: 10.1109/ipact.2017.8245212. [Online]. Available: https://doi.org/10.1109%2Fipact.2017.8245212.

[97] R. Vilimek and T. O. Zander, "BC(eye): Combining eye-gaze input with brain-computer interaction", in *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, Springer Berlin Heidelberg, 2009, pp. 593–602. DOI: 10.1007/978-3-642-02710-9_66. [Online]. Available: https://doi.org/10.1007%2F978-3-642-02710-9_66.

[98] E. C. Lee, J. C. Woo, J. H. Kim, M. Whang, and K. R. Park, "A brain–computer interface method combined with eye tracking for 3d interaction", *Journal of Neuroscience Methods*, vol. 190, no. 2, pp. 289–298, Jul. 2010. DOI: 10.1016/j.jneumeth.2010.05.008. [Online]. Available: https://doi.org/10.1016%2Fj.jneumeth.2010.05.008.

[99] A. Petrushin, J. Tessadori, G. Barresi, and L. S. Mattos, "Effect of a click-like feedback on motor imagery in EEG-BCI and eye-tracking hybrid control for telepresence", in *2018 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, IEEE, Jul. 2018. DOI: 10.1109/aim.2018.8452292. [Online]. Available: https://doi.org/10.1109%2Faim.2018.8452292.

[100] B. H. Kim, M. Kim, and S. Jo, "Quadcopter flight control using a low-cost hybrid interface with EEG-based classification and eye tracking", *Computers in Biology and Medicine*, vol. 51, pp. 82–92, Aug. 2014. DOI: 10.1016/j.compbiomed.2014.04.020. [Online]. Available: https://doi.org/10.1016%2Fj.compbiomed.2014.04.020.

[101] P. Stawicki, F. Gembler, A. Rezeika, and I. Volosyak, "A novel hybrid mental spelling application based on eye tracking and SSVEP-based BCI", *Brain Sciences*, vol. 7, no. 12, p. 35, Apr. 2017. DOI: 10.3390/brainsci7040035. [Online]. Available: https://doi.org/10.3390%2Fbrainsci7040035.

[102] T. Malechka, T. Tetzel, U. Krebs, D. Feuser, and A. Graeser, "sBCI-headset—wearable and modular device for hybrid brain-computer interface", *Micromachines*, vol. 6, no. 3, pp. 291–311, Feb. 2015. DOI: 10.3390/mi6030291. [Online]. Available: https://doi.org/10.3390%2Fmi6030291.

[103] R. P. Rao, *Brain-computer interfacing: An introduction.* Cambridge University Press, 2013.

[104] J. G. Betts, *Anatomy and physiology.* OpenStax College, Rice University, 2013.

[105] W. J. Germann, C. L. Stanfield, and L. Fadiga, *Fisiologia umana.* Edises, 2004.

[106] P. A. Iaizzo, "Introduction to neurophysiology", in *Neural Engineering*, B. He, Ed. Boston, MA: Springer US, 2013, pp. 1–86, ISBN: 978-1-4614-5227-0. DOI: 10.1007/978-1-4614-5227-0_1. [Online]. Available: https://doi.org/10.1007/978-1-4614-5227-0_1.

[107] S. Cristina and K. P. Camilleri, "Unobtrusive and pervasive video-based eye-gaze tracking", *Image and Vision Computing*, vol. 74, pp. 21–40, Jun. 2018. DOI: 10.1016/j.imavis.2018.04.002. [Online]. Available: https://doi.org/10.1016%2Fj.imavis.2018.04.002.

[108] http://www.gtec.at, [Online; accessed 08-January-2019].

[109] http://www.theeyetribe.com, [Online; accessed 10-January-2019].

[110] H. He, Y. Bai, E. A. Garcia, and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning", in *Neural Networks, 2008. IJCNN 2008.(IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, IEEE, 2008, pp. 1322–1328.

[111] M. Schreuder, B. Blankertz, and M. Tangermann, "A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue", *PloS one*, vol. 5, no. 4, e9813, 2010.

[112] F. Ferracuti, A. Freddi, S. Iarlori, S. Longhi, and P. Peretti, "Auditory paradigm for a p300 bci system using spatial hearing", in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, IEEE, 2013, pp. 871–876.

[113] S. Halder, K. Takano, H. Ora, A. Onishi, K. Utsumi, and K. Kansaku, "An evaluation of training with an auditory p300 brain-computer interface for the japanese hiragana syllabary", *Frontiers in neuroscience*, vol. 10, p. 446, 2016.

[114] M. Huang, J. Jin, Y. Zhang, D. Hu, and X. Wang, "Usage of drip drops as stimuli in an auditory p300 bci paradigm", *Cognitive neurodynamics*, vol. 12, no. 1, pp. 85–94, 2018.

[115] M. Schreuder, T. Rost, and M. Tangermann, "Listen, you are writing! speeding up online spelling with a dynamic auditory bci", *Frontiers in neuroscience*, vol. 5, p. 112, 2011.

[116] M. Guo, G. Xu, L. Wang, and J. Wang, "Research on auditory bci based on wavelet transform", in *Virtual Environments Human-Computer Interfaces and Measurement Systems (VECIMS), 2012 IEEE International Conference on*, IEEE, 2012, pp. 171–175.

[117] S. Halder, E. M. Hammer, S. C. Kleih, M. Bogdan, W. Rosenstiel, N. Birbaumer, and A. Kübler, "Prediction of auditory and visual p300 brain-computer interface aptitude", *PloS one*, vol. 8, no. 2, e53513, 2013.

[118] E. M. Hammer, S. Halder, S. C. Kleih, and A. Kübler, "Psychological predictors of visual and auditory p300 brain-computer interface performance", *Frontiers in neuroscience*, vol. 12, 2018.