Enhancing ERP-based Brain-Computer Interfaces for Practical Applications

Asynchrony, Deep Learning and a Novel BCI Platform

Doctoral Dissertation

Eduardo Santamaría-Vázquez

Advisor: Roberto Hornero

Universidad de Valladolid

Doctoral Program of
Information and Telecommunication Technologies

Doctoral Dissertation

Enhancing ERP-based Brain-Computer Interfaces for Practical Applications: Asynchrony, Deep Learning, and a Novel BCI Platform

Dissertation presented by **D. Eduardo Santamaría-Vázquez**
in partial fulfillment of the requirements for the *Ph.D.* Degree
at the *University of Valladolid*

Directed by:
Dr. Roberto Hornero Sánchez

2023
Valladolid, Spain

El cerebro es un mundo que consta de numerosos continentes inexplorados y grandes extensiones de territorios desconocidos.
Santiago Ramon y Cajal

Defense

Título	Enhancing ERP-based Brain-Computer Interfaces for Practical Applications: Asynchrony, Deep Learning, and a Novel BCI Platform
Autor	D. Eduardo SantamaríaVázquez
Director	Dr. D. Roberto Hornero Sánchez
Departamento	Teoría de la Señal y Comunicaciones e Ingeniería Telemática

---

Tribunal

Presidente

Vocal

Secretario

acuerda otorgarle la calificación de

En Valladolid, a de del

Universidad de Valladolid

Escuela Técnica Superior de Ingenieros de Telecomunicación
Dpto. de Teoría de la Señal y Comunicaciones e Ingeniería Telemática

Research Stay for the International Mention

City:	Boston (United States of America)
Faculty:	Harvard Medical School
Institution:	Spaulding Research Institute
Research group:	Spaulding Neuromodulation Center
Dates:	01/07/2022-10/10/2022
Duration:	3 months and 10 days
Supervisor:	Dr. Leon Morales-Quezada

SPAULDING RESEARCH INSTITUTE

HARVARD UNIVERSITY

Agradecimientos

En primer lugar, quisiera expresar mi más sincera gratitud al Dr. Roberto Hornero Sánchez por guiarme durante estos años. Gracias a la oportunidad que me brindaste he encontrado una vocación que disfruto cada día; y gracias a tus consejos, compromiso, sinceridad y cercanía no solo me he desarrollado enormemente a nivel académico y profesional, sino también como persona.

Quiero dar las gracias también al resto de miembros del Grupo de Ingeniería Biomédica (GIB), a los que están y a los que estuvieron, por hacer de esta etapa un tiempo inolvidable. El GIB no solo es un grupo de investigación de primer nivel, sino también una familia de personas excepcionales, de esas que generan un ambiente estupendo, te ayudan y hacen que crezcas cada día. Las ideas locas en los cafés y '*happy meals*', los congresos, las quedadas y los viajes han hecho que esta etapa haya sido inolvidable. Mención especial para Víctor, por todos los momentos que hemos compartido juntos: la tensión de las demostraciones, las innumerables discusiones en el laboratorio para poner ideas en común, la decepción al encontrar otro *bug* más, o la emoción de cuando todo funcionaba. Sin tí, este camino habría sido mucho más dificil.

I would also like to thank Dr. Leon Morales Quezada and the Spaulding Neuromodulation Center for embracing me at this internationally renowned institution. This experience was unforgettable for many reasons, and it has allowed me to grow professionally and personally. I would like to express my gratitude to Anayali and Ine for their invaluable help; as well as Raul, Logan, Daniel, Nico, and the rest of the people at center for making me feel at home.

Por último, quiero expresar mi más profundo agradecimiento a mis padres, mi hermano y Cristina. Gracias por vuestro constante apoyo, por hacer que sea mejor persona cada día, y por inculcarme unos principios de los que me enorgullezco. Sin duda, esto no habría sido posible sin vosotros.

Abstract

Throughout history, humans have sought ways to break free from the constraints of the body and interact with the world directly through the mind. Brain-computer interfaces (BCIs) represent the realization of this long-standing ambition, allowing individuals to control external devices directly using their brain activity. BCIs measure brain activity using electroencephalography (EEG), a technique that records the electrical activity of neurons using electrodes placed over the scalp. Then, EEG is analyzed using signal processing methods to decode users' intentions and translate them into commands that can be used to control external devices.

BCI technology has great potential in a wide range of applications within the biomedical sector. Its most direct use is the development of assistive systems for people with severe motor impairments or paralysis, allowing them to communicate and control devices. Moreover, the implementation of BCIs in the neurorehabilitation field is rapidly increasing in recent years. Currently, BCIs are used as an innovative technique to treat various conditions, including stroke, depression, drug addiction, and even to enhance human cognitive abilities, such as memory or attention. However, there are several barriers to wider adoption of BCIs, including low reliability, their synchronous operation, and a lack of validation with target users. In order to fully realize the potential of BCIs, these limitations must be addressed.

The research presented in this compendium of publications was focused on assistive systems based on event-related potentials (ERPs), which use the P300 evoked potential as control signal. The aim of this work was to address some of the current limitations of these BCIs in assistive contexts by developing new signal processing methods and software tools to accelerate research and applications in this area.

Firstly, the issue of the asynchronous control was addressed. By default, these

systems have a synchronous behavior: they select a command even when the user is not attending the stimuli. This poses a great limitation that prevents their use for practical applications outside the laboratory. For this reason, we proposed a novel method, called the Oddball Steady State Response Detection (OSRD) method, for monitoring the user's attention and avoiding undesired selections when the user is engaged in another task. This method aimed to discriminate the user's control state by detecting the steady-state visual evoked potential (SSVEP) provoked by the stimulation pattern of the row-column paradigm (RCP) used in our ERP-based BCI. After a comprehensive characterization experiment with five subjects, we concluded that this control signal could be used to detect the user's control state in assistive ERP-based spellers. Then, the system was designed using two stages: the control state detection stage, which monitors the user's attention and provides asynchronous control with the OSRD method, and the command decoding stage, which decodes the desired command only if the user is attending to the stimulation paradigm. The OSRD method uses a feature engineering approach based on spectral and correlation features. It was tested on 15 subjects in offline and online sessions. The results showed that our approach provided reliable asynchronous control, with an final average accuracy of 95.5% in the control state detection stage during the online sessions. Additionally, it has two advantages in comparison to previous methods based on thresholds: it is independent of the command decoding stage, and it does not require to extend the duration of the calibration sessions.

Once this first asynchronous approach was designed, we investigated the use of deep learning to improve the command decoding accuracy and speed of our system. A novel convolutional neural network (CNN), called EEG-Inception, was proposed. This model was inspired in different state-of-the-art methods from the computer vision domain to increase its performance. Concretely, this lightweight CNN was designed to perform a multiscale analysis of the EEG signal using Inception modules, which were combined with a hyperparamter optimization process and different techniques to avoid overfitting. Additionally, this architecture was trained with a novel strategy based on cross-subject transfer learning and fine-tuning to improve the performance of the model in the command decoding task with minimum calibration time. The model was validated with data from 73 subjects, including 31 with severe motor disabilities, for the command decoding task. The results showed that EEG-Inception outperformed five previous approaches in a direct comparison, achieving significant accuracy improvements up to 16.0%, 10.7%, 7.2%, 5.7% and 5.1% in comparison to regularized discriminant analysis,

xDAWN with Riemannian geometry, CNN-BLSTM, DeepConvNet and EEGNet, respectively.

The next study addressed the applicability of EEG-Inception to detect the user's control state. We designed a signal processing pipeline that leveraged the power of this deep-learning model to improve the asynchronous control of our system. It is worth noting that this was the first asynchronous ERP-based BCI reported in the literature that solely relied on deep learning for both the control state detection and command decoding stages. This approach method was evaluated with 22 healthy subjects. The results for the control state detection task indicated that the model achieved an average test accuracy of over 91% with just one sequence of stimulation and 30 calibration trials, reaching a peak value of 96.95% when using 15 sequences. Furthermore, the overall system, including the control state detection and command decoding stages, was able to achieve a high information transfer rate of 35.54 bpm. These results demonstrate that the proposed system is a promising step towards more practical applications of these BCIs.

Finally, we created MEDUSA©, a novel software ecosystem optimized for BCI and cognitive neuroscience research. Our goal was to overcome the limitations of existing BCI platforms (e.g., BCI2000, OpenVibe) in terms of modularity, flexibility and scalability. MEDUSA© was implemented in Python, a popular open-source programming language that is commonly used in both research and industry. The reasons of this popularity are its high-level of abstraction, which reduces the development time, and strong community support, with a wide range of packages and specialized tools for signal processing and machine learning, such as Numpy, Scipy or Tensorflow. MEDUSA© has two main components: (1) a Python library called MEDUSA© Kernel, which provides signal processing functions at different levels of abstraction; and (2) a BCI platform called MEDUSA© Platform, which implements advanced signal acquisition based on the lab-streaming layer protocol, a wide variety of real-time charts, and open implementations of state-of-the-art BCI and neuroscience experiments, including our ERP-based BCI based on deep learning. Additionally, this open-source software encourage community contributions in order to increase its impact and provide a space to share the latest advances in the field.

We believe that the findings and developments presented in this doctoral dissertation will enhance the practical use of ERP-based BCIs outside the laboratory, particularly as assistive technology to improve the lives of individuals with disabilities. Our work highlighted that both the control state detection and com-

mand decoding are equally important for assistive systems, proposing novel signal processing methods for both stages. In this regard, the final design of our asynchronous ERP-based BCI, the first that fully relies on deep learning, represents a significant step forward in this context. Additionally, MEDUSA© is a novel software ecosystem that meets the needs of BCI researchers by providing a range of signal processing functions, deep-learning architectures, connectivity analysis tools, and ready-to-use BCI and neuroscience experiments. Thus, it has the potential to accelerate BCI and cognitive neuroscience research.

Acronyms

ADHD	Attention-Deficit/Hyperactivity Disorder
BCI	Brain–Computer Interface
BE	Backward Elimination
BH	Benjamini-Hochberg correction
BOLD	Blood-Oxygen-Level-Dependent response
CAR	Common Average Reference
CCA	Canonical Correlation Analysis
CNS	Central Nervous System
C-VEP	Code-Modulated Visual Evoked Potential
EEG	Electroencephalography
ERD	Event-Related Desynchronization
ERP	Event-Related Potential
ERS	Event-Related Synchronization
FDR	False Discovery Rate
FIR	Finite Impulse Response
fMRI	Functional Magnetic Resonance Imaging
fNIRS	Functional Near-Infrared Spectroscopy
HS	Healthy Subjects
IIR	Infinite Impulse Response
ISI	Inter-Stimuli Interval
JCR	Journal Citation Reports
LIS	Locked-In Syndrome
LDA	Linear Discriminant Analysis
LOO	Leave one out
MD	Motor-Disabled Subjects
MEG	Magnetoencephalography

MI	Motor Imagery
MSE	Multiscale Entropy
NF	Neurofeedback
OSRD	Oddball Steady State Response Detection
PET	Positron Emission Tomography
PSD	Power Spectral Density
RCP	Row-Col Paradigm
RG	Riemannian geometry
ROC	Receiving Operating Characteristic
RSVP	Rapid Serial Visual Presentation
SampEn	Sample Entropy
SCI	Spinal Cord Injury
SCP	Slow Cortical Potentials
SD	Stimulus Duration
SMR	Sensorimotor Rhythms
SOA	Stimulus Onset Asynchrony
SSVEP	Steady-State Visual Evoked Potential
SVM	Support Vector Machine
SW	Step-Wise regression
VEP	Visual Evoked Potential

Contents

Abstract **I**

Acronyms **V**

1 Introduction **1**

1.1 Compendium of publications: thematic consistency 2

1.2 Dissertation Context 7

1.3 Brain–computer interfaces 9

1.3.1 Recording neural activity 10

1.3.2 EEG 18

1.3.3 BCI control signals 20

1.3.4 EEG processing in BCI 22

1.3.5 BCI Applications 24

1.4 State of the art 25

1.4.1 Asynchronous control of ERP-based BCIs 25

1.4.2 Deep learning in ERP-based BCIs 27

1.4.3 BCI platforms 29

2 Hypothesis and objectives **33**

2.1 Hypothesis 33

2.2 Objectives 34

3 Subjects and signals **37**

3.1 The row-column paradigm 37

3.2 Subjects 39

3.3 EEG recording setup 41

4 Methods **45**
4.1 Pre-processing . . . 45
4.1.1 Frequency filtering . . . 46
4.1.2 Spatial filtering . . . 47
4.2 Feature extraction . . . 47
4.2.1 Time-based features . . . 48
4.2.2 Frequency-based features . . . 48
4.2.3 Correlation-based features . . . 49
4.3 Feature selection . . . 51
4.3.1 Stepwise regression . . . 51
4.4 Feature classification . . . 52
4.4.1 Linear Discriminant Analysis . . . 53
4.4.2 EEG-Inception . . . 54
4.5 Validation . . . 59
4.5.1 Performance metrics . . . 59
4.5.2 Statistical analysis . . . 60
4.5.3 Cross-validation . . . 61
4.6 BCI platform development . . . 62
4.6.1 Components . . . 62
4.6.2 Design principles . . . 62
4.6.3 Implemented in Python . . . 63
4.6.4 Open-source community-oriented philosophy . . . 64

5 Results **65**
5.1 SSVEPs elicited by the RCP . . . 65
5.2 Oddball Steady Response Detection (OSRD) method . . . 67
5.2.1 Offline experiment . . . 69
5.2.2 Online experiment . . . 70
5.3 Command decoding with EEG-Inception . . . 72
5.3.1 Hyperparameter optimization . . . 73
5.3.2 Evaluation experiment . . . 74
5.4 Control state detection with EEG-Inception . . . 75
5.4.1 Evaluation experiment . . . 78
5.5 MEDUSA©: our novel BCI platform . . . 82
5.5.1 MEDUSA© Kernel . . . 82
5.5.2 MEDUSA© Platform . . . 84

6 Discussion **91**
6.1 SSVEP elicited by non-target stimuli from ERP-based spellers . . 92
6.2 Feature-engineering approach for control state detection: the OSRD method 94
6.2.1 Evaluation experiment 94
6.2.2 Comparison with previous approaches 95
6.3 Deep-learning approach for command decoding: EEG-Inception . . 98
6.3.1 Architecture design 98
6.3.2 Evaluation experiment 99
6.4 Deep-learning approach for control state detection: EEG-Inception 101
6.4.1 Evaluation experiment 102
6.4.2 Comparison with previous approaches 103
6.5 Design of MEDUSA© and comparison with previous BCI platforms 105
6.6 Limitations 109

7 Conclusions **111**
7.1 Contributions 111
7.2 Main conclusions 113
7.3 Future research lines 114

A Compendium of publications **117**

B Scientific achievements **119**
B.1 Publications 119
B.1.1 Papers indexed in the JCR 119
B.1.2 Book chapters 121
B.1.3 International conferences 121
B.1.4 National conferences 123
B.2 International internship 125
B.3 Awards and honors 127

C Resumen en castellano **129**
C.1 Introducción 129
C.2 Hipótesis y objetivos 133
C.3 Sujetos 135
C.4 Métodos 137
C.5 Resultados y discusión 139
C.6 Conclusiones 143

Bibliography **145**

Index **155**

List of Figures

1.1 Main contributions of the papers included in the compendium of publications, arranged by the BCI function addressed in each study and ordered chronologically. IEEE TNSRE: IEEE Transactions on Neural Systems and Rehabilitation Engineering; CMPB: Computer Methods and Programs in Biomedicine; BCI: brain-computer interface; EEG: electroencephalography; ERP: event-related potentials; RCP: row-column paradigm; SSVEP: steady-state visual evoked potentials. 3

1.2 Schematic representation of a EEG-based BCI system. Firstly, the signal is acquired with electrodes over the scalp. Then, the signal processing stage decodes the user's intentions and translated them into commands. Finally, the commands are sent to the final application, which provides feedback to the user. BCI applications can replace, restore, enhance, supplement and/or improve the natural body outputs. Adapted with permission from Martínez-Cagigal (2020) 10

1.3 Schematic representation of a neuron. Dendrites receive electrical impulses from other neurons. The cell body, or soma, is the core section of the neuron, maintaining its structure and providing energy to drive function. The axon transmits electrical impulses to the axon terminals, and can be coated in myelin, a protein that improves the transmission. Dendrites and axon terminals are connected by synapses. Adapted from He (2020). 11

1.4 Schematic representation of the brain, divided in its three anatomical parts: brainstem, cerebellum and cerebrum. The lobes of the cerebral cortex are also displayed. Adapted from Jacobson (2008). 13

1.5 The International System 10-10 illustrated in schematic views showing (A) a lateral view, (B) a top view, and (C) a scalp projection. This standard uses proportional distances of 10% between the nasion and the inion for positioning the electrodes on the scalp (Martínez-Cagigal, 2020). 20

1.6 The figure illustrates the averaged EEG epochs for both target and non-target stimuli in an ERP-based BCI. The left graph represents the signal in channel Cz, with the stimulus onset placed at t=0. As can be seen, the ERP, with the P300 component, is clearly visible for target stimuli. The graph on the right shows the amplitude values recorded in eight different channels at the peak of the ERP. 23

3.1 Representation a row-column paradigm (RCP) trial with a 6×6 matrix. When the trial starts, the system highlights all the rows and columns of the matrix. The process is repeated N_s times to increase the number of observations and improve the detection accuracy of the P300 potential. Finally, the BCI processes the EEG signal and decodes the command. 38

3.2 Electrode montages according to the International System 10–10: (A) Asynchrony database (Santamaría-Vázquez et al., 2019a, 2022); (B) BCI Web Browser and BCI Social Networks databases (Santamaría-Vázquez et al., 2020b). Ground and reference electrodes are marked in yellow and blue, respectively. Adapted from Martínez-Cagigal (2020). 42

4.1 (a) Representation of the frequency-based feature, including the parameters f_{st}, bw_1, bw_2. The graph shows the averaged power spectral density (PSD) of control and non-control trials of the Asynchrony database. (b) Schematic representation of the canonical correlation analysis (CCA) used to extract the correlation-based. The EEG signal represents one trial of dimensions $C \times N$, where C is the number of channels and N the trial length in samples. The reference signal is an ideal sinusoid $sin(2\pi \cdot f_{st} \cdot t)$ of dimensions $1 \times N$. 50

4.2 Schematic representation of EEG-Inception. 2D convolution blocks and depthwise 2D convolution blocks include batch normalization, non-linear activation and dropout regularization. The kernel size is displayed for convolutional and average pooling layers. 56

5.1 Matrices used during the experiments: (a) RCP matrix, (b) overt matrix, (c) covert matrix. 66

5.2 SSVEPs for different stimulation frequencies. Graphs on the left show the grand average of the PSD for all trials, channels and participants. Topographic plots on the right show the normalized peak value of the PSD at the stimulation frequency averaged for all trials and participants. 68

5.3 Temporal and spectral representation of the ERPs in overt, covert, control and non-control modes for participant S01. The upper part of the figure shows the averaged epochs (i.e., 1000 ms after stimuli) in channel Cz. The shaded area represents the 95% confidence interval. The bottom part of the figure shows the averaged PSD in channel Cz. The stimulation frequency in overt and control modes was 5.71 Hz. 68

5.4 Experimental setup. (a) Depicts a schematic representation of the subject and the two screens. The left screen displayed the browser used during the non-control runs while the speller was shown on the right screen. The paradigm was active during both tasks, but subjects were only required to pay attention to the stimuli during the control runs. (b) Shows the experiment plan, which was made up of two sessions, each with 10 runs of 6 trials and 15 sequences. Control and non-control runs were alternated to prevent subject fatigue. 71

5.5 Workflow of the study. First, the hyperparameter optimization process for EEG-Inception was done based using the validation set. Then, the models were trained using the training set. Finally, they were evaluated in the test set using a fine-tuning process for each subject. Fine-tuning trials were randomly selected, and the process is repeated 100 times for each N and subject. 73

5.6 Results of the optimization process of EEG-Inception in the validation set (8 healthy subjects). Each line depicts the mean command decoding accuracy after a fine-tuning process for each subject with $N = 30$, considering 5 sequences of stimulation. (a) Score for each activation function and dropout rate using the best value of learning rate ($lr = 0.001$). (b) Score for each learning rate and dropout rate using the best activation function (ELU). 74

5.7 ITR: information transfer rate; N: number of fine-tuning trials. ITR in bits/min in the test set (31 motor disabled subjects) as a function of N and the number of sequences. 75
5.8 Normalized confusion matrices averaged across subjects. 80
5.9 n.u: normalized units. Characterization of correctly and incorrectly classified trials for the control state detection task. The upper graphs show the averaged EEG epochs for the 3 different conditions: non-control, control non-target and control target. The lower graphs show the power spectral density of the entire trials. 81
5.10 ITR: information transfer rate (bpm); N: number of fine-tuning trials in control state for each subject. Average ITR of the overall system, including control state detection and command decoding stages and only considering control trials. 81
5.11 Schematic overview of MEDUSA©. EEG: electroencephalography; ECG: electrocardiography; EMG: electromyography; LSL: lab-streaming layer. An arbitrary number of input signals can be received in the platform through the LSL. These signals are available for real-time charts and apps, which implement open-loop or closed-loop BCI and neuroscience experiments. In this case, we represent the row-column paradigm (RCP) app. Functions from the Kernel may be used for signal processing in real-time charts and apps. In this example, the model detects the event-related potentials (ERP) in the EEG. 84
5.12 Main window of MEDUSA© Platform. The view is divided in panels to control the different functionalities. These panels are: apps (up-left), log (bottom-left) and real-time plots (right). Additionally, more controls and configurations are available in the task bars. . . 85
5.13 Screenshots of some MEDUSA© apps. (a) row-column paradigm (RCP) speller, (b) code-modulated visual evoked potentials (c-VEP) speller, (c) Motor imagery paradigm, (d) Jedi cube scenario from the neurofeedback app, (e) Luke's spaceship scenario from the neurofeedback app, (f) Neurorunner scenario from the neurofeedback app, (g) Digit Span test from the Neuropsychological evaluation app (h) Corsi Block-Tapping test from the Neuropsychological evaluation app, (i) Dual N-Back test from the Neuropsychological evaluation app. 87

List of Tables

1.1 Previous deep-learning approaches for ERP-based BCIs 29

3.1 Demographics of the Asynchrony database. 41
3.2 Demographics and clinical data of the BCI Web Browser database. . . . 42
3.3 Demographics and clinical data of the BCI Social Networks database. . 43
3.4 Summary of each database. 43

4.1 EEG-Inception architecture details 58

5.1 Stimulation parameters for the characterization experiment. 67
5.2 Results of the offline experiment 71
5.3 Results of the online experiment 72
5.4 Command decoding accuracy 76
5.5 Control state detection accuracy using EEG-Inception 79
5.6 Overall accuracy of the asynchronous system using EEG-Inception . . . 80

6.1 Comparative of the SSVEP power and speed of selection 93
6.2 Comparative of previous asynchronous ERP-based spellers. 96
6.3 Comparative of previous asynchronous systems with independent control state detection and command decoding pipelines. 104
6.4 Comparative between different platforms for BCI experiments. . . 108

C.1 Resumen de las bases de datos utilizadas. 136

Chapter 1

Introduction

This doctoral dissertation aims to design, develop, and test novel methods and tools for the advancement of non-invasive brain–computer interfaces (BCI) with one objective: improve their applicability in practical environments. To this end, we investigated several crucial aspects of these systems, such as asynchronous control, classification performance and BCI platforms. This research has led to the publication of 4 articles in journals indexed in the Journal Citation Reports (JCR) from the Web of Science™. This scientific production has led to write this dissertation as a compendium of publications. Throughout the document, a general framework of the dissertation is provided to summarize and discuss the most relevant results and draw a joint conclusion of the study.

The structure of this chapter is as follows. First, the thematic consistency of the articles that make up this dissertation is explained in detail in section 1.1. Then, the context of this research and the related knowledge fields are introduced in section 1.2. Section 1.3 is focused on BCIs, introducing the history, characteristics, control signals, applications, and limitations of these systems, with the latter giving rise to the motivation for the present study. In this section, several non-invasive methods to record brain activity are also discussed to motivate the use of electroencephalography (EEG) in this work, explaining the main advantages and disadvantages of this biomedical signal. Finally, section 1.4 provides a comprehensive state-of-the-art revision of the main topics that have been addressed in this investigation: asynchronous control, deep learning for EEG decoding, and BCI platforms.

1.1 Compendium of publications: thematic consistency

Understanding the brain has been a long-standing ambition for humankind, but this task proved to be a real challenge during centuries. Until the late nineteenth and early twentieth centuries, when Santiago Ramón y Cajal and other pioneers laid the groundwork of contemporary neuroanatomy, the rate of developments in this field was slow (Fornito et al., 2016). Nowadays, recent advances in brain imaging opened a window to our central nervous system that was impossible just a few decades ago, transforming the field of neuroscience into one of the most fruitful research areas. In this context, both research laboratories and industry are exploiting this knowledge to build BCIs, a technology with the potential to transform fields such as human-machine interaction, neurorehabilitation, and entertainment, among others (Wolpaw and Wolpaw, 2012). This technology analyzes the EEG, which measures the electrical activity of the brain, to decode user's intentions in real time, allowing direct interaction with devices without the use of muscles or peripheral nerves. Nevertheless, there are still limitations that need to be addressed in order to use these systems for practical day-to-day applications (Wolpaw and Wolpaw, 2012). The most important are the complexity of the required hardware, the difficulty of fine measurement of brain activity, limited performance, synchronous operation and lack of validation with end users (Santamaría-Vázquez et al., 2019a).

This dissertation aims to contribute with novel signal processing methods and tools to improve the usability of BCIs for practical applications. The thread running through this research is the use of BCIs based on event-related potentials (ERP) as a novel assistive technology to improve the independence and quality of life of severely disabled people. Thus, the four articles that comprise the compendium of publications are focused on this field. A graphical representation of the thematic consistency and contributions of the four papers is depicted in Figure 1.1.

The first paper was focused on the problem of asynchronous control of ERP-based BCIs (Santamaría-Vázquez et al., 2019a). By default, these systems decide which command the user wants every few seconds. However, current implementations cannot detect whether the user is controlling the system or is engaged in other task, performing synchronous selections even when it is not desired. Thus, an additional method is required in order to detect the user's attention focus, a task that has turned out to be as difficult as ERP detection. This first study proposed a novel feature-engineering method based on the power spectral density (PSD) and

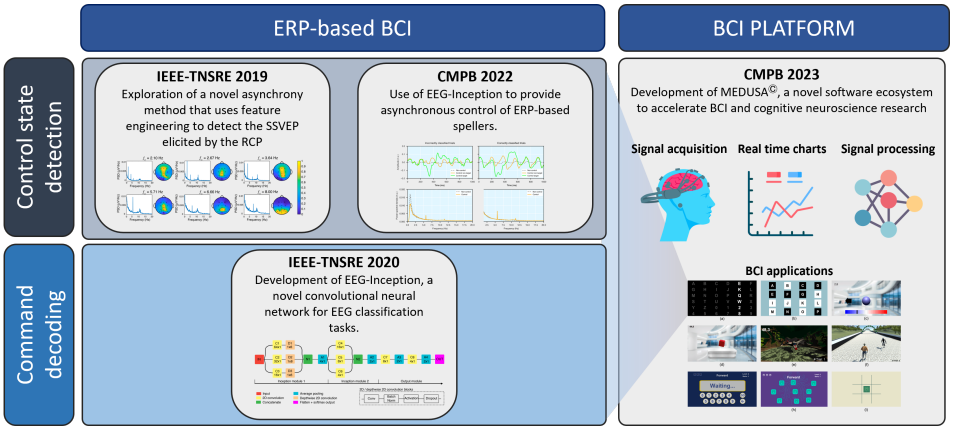

Figure 1.1: Main contributions of the papers included in the compendium of publications, arranged by the BCI function addressed in each study and ordered chronologically. IEEE TNSRE: IEEE Transactions on Neural Systems and Rehabilitation Engineering; CMPB: Computer Methods and Programs in Biomedicine; BCI: brain-computer interface; EEG: electroencephalography; ERP: event-related potentials; RCP: row-column paradigm; SSVEP: steady-state visual evoked potentials.

canonical correlation analysis (CCA) to investigate the potential of residual steady state visual evoked potentials (SSVEP) to provide a robust asynchronous control of ERP-based BCIs. This approach was tested with 15 healthy subjects, showing its advantages. The second article was aimed to improve the classification performance in this BCI paradigm (Santamaría-Vázquez et al., 2020b). To this end, we investigated the advantages of convolutional neural networks (CNN) for EEG processing. Concretely, we proposed a novel architecture, called EEG-Inception, which was able to overcome previous approaches for ERP detection in a population of 73 subjects, including 31 motor disabled. In the third article, we returned to the limitation of synchronous operation in ERP-based BCIs (Santamaría-Vázquez et al., 2022). This time, we investigated a novel strategy for the asynchronous control task using EEG-Inception. This strategy was tested in 22 healthy subjects, achieving significantly higher performance than our previous method. Finally, in the fourth article we presented MEDUSA©, a novel open-source, Python-based, software ecosystem to facilitate the implementation of BCI and cognitive neuroscience experiments with the objective of accelerating research in these areas (Santamaría-Vázquez et al., 2023). This platform was developed during the whole dissertation, and it implements the methods from previous articles to foster its use and reproducibility. It also includes a wide range of signal processing algorithms and BCI applications that were implemented for side projects developed in parallel

to this research, making MEDUSA© one of the most complete solutions available today.

The four articles that make up this dissertation, which is organized as a compendium of publications, are included in this document in Appendix A. Titles, authors, and abstracts of each one, as well as the journals in which they were published, are shown below:

Asynchronous Control of ERP-based BCI Spellers Using Steady-State Visual Evoked Potentials Elicited by Peripheral Stimuli (Santamaría-Vázquez et al., 2019a).

Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Javier Gomez-Pilar, and Roberto Hornero. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27(9), p. 1883-1892, 2019. Impact factor in 2019: 3.340, D1 (Q1) in "Rehabilitation", and Q1 in Biomedical Engineering (JCR-WOS).

Abstract: Brain–computer interface (BCI) spellers based on event related potentials (ERPs) are intrinsically synchronous systems. Therefore, selections are constantly made, even when users are not paying attention to the stimuli. This poses a major limitation in real-life applications, in which an asynchronous control is required. The aim of this study is to design, develop and test a novel method to discriminate whether the user is controlling the system (i.e., control state) or is engaged in other task (i.e., non-control state). To achieve such asynchronous control, our method detects the steady-state visual evoked potentials (SSVEPs) elicited by peripheral stimuli of ERP-based spellers. A characterization experiment was conducted to investigate several aspects of this phenomenon. Then, the proposed method was validated in offline and online sessions. A total of 20 healthy subjects participated the experiments. The proposed method achieved an average accuracy of 95.5% for control state detection during the online sessions, providing a reliable asynchronous control. Furthermore, our approach is independent of the ERP classification stage, and to the best of our knowledge, is the first procedure that does not need to extend the duration of the calibration sessions to acquire non-control observations.

EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-based Brain-Computer Interfaces (Santamaría-Vázquez et al., 2020b).

Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Fernando Vaquerizo-Villar, and Roberto Hornero. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28(12), p. 2773-2782, 2020. Impact factor in 2020: 3.802, D1 (Q1) in "Rehabilitation" and Q2 in Biomedical Engineering (JCR-WOS).

Abstract: In recent years, deep-learning models gained attention for electroencephalography (EEG) classification tasks due to their excellent performance and ability to extract complex features from raw data. In particular, convolutional neural networks (CNN) showed adequate results in brain-computer interfaces (BCI) based on different control signals, including event-related potentials (ERP). In this study, we propose a novel CNN, called EEG-Inception, which improves the accuracy and calibration time of assistive ERP-based BCIs. To the best of our knowledge, EEG-Inception is the first model to integrate Inception modules for ERP detection, which combined efficiently with other structures in a light architecture, improved the performance of our approach. The model was validated in a population of 73 subjects, of which 31 present motor disabilities. Results show that EEG-Inception outperforms 5 previous approaches, yielding significant improvements for command decoding accuracy up to 16.0%, 10.7%, 7.2%, 5.7% and 5.1% in comparison to rLDA, xDAWN + Riemannian geometry, CNN-BLSTM, DeepConvNet and EEGNet, respectively. Moreover, EEG-Inception requires very few calibration trials to achieve state-of-the-art performances taking advantage of a novel training strategy that combines cross-subject transfer learning and fine-tuning to increase the feasibility of this approach for practical use in assistive applications.

Robust Asynchronous Control of ERP-Based Brain-Computer Interfaces using Deep Learning (Santamaría-Vázquez et al., 2022).

Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Sergio Pérez-Velasco, Diego Marcos-Martínez, and Roberto Hornero. *Computer Methods and Programs in Biomedicine*, vol. 215, 2022. Impact factor in 2021: 7.027, Q1 in "Computer Science, Interdisciplinary Applications", "Computer Science, Theory & Methods", "Biomedical Engineering" and "Medical Informatics" (JCR-WOS).

Abstract: Background and Objective. Brain-computer interfaces (BCI) based on event-related potentials (ERP) are a promising technology for alternative and augmented communication in an assistive context. However, most approaches

to date are synchronous, requiring the intervention of a supervisor when the user wishes to turn his attention away from the BCI system. In order to bring these BCIs into real-life applications, a robust asynchronous control of the system is required through monitoring of user attention. Despite the great importance of this limitation, which prevents the deployment of these systems outside the laboratory, it is often overlooked in research articles. This study was aimed to propose a novel method to solve this problem, taking advantage of deep learning for the first time in this context to overcome the limitations of previous strategies based on hand-crafted features. *Methods.* The proposed method, based on EEG-Inception, a novel deep convolutional neural network, divides the problem in 2 stages to achieve the asynchronous control: (*i*) the model detects user's control state, and (*ii*) decodes the command only if the user is attending to the stimuli. Additionally, we used transfer learning to reduce the calibration time, even exploring a calibration-less approach. *Results.* Our method was evaluated with 22 healthy subjects, analyzing the impact of the calibration time and number of stimulation sequences on the system's performance. For the control state detection stage, we report average accuracies above 91% using only 1 sequence of stimulation and 30 calibration trials, reaching a maximum of 96.95% with 15 sequences. Moreover, our calibration-less approach also achieved suitable results, with a maximum accuracy of 89.36%, showing the benefits of transfer learning. As for the overall asynchronous system, which includes both stages, the maximum information transfer rate was 35.54 bpm, a suitable value for high-speed communication. *Conclusions.* The proposed strategy achieved higher performance with less calibration trials and stimulation sequences than former approaches, representing a promising step forward that paves the way for more practical applications of ERP-based spellers.

MEDUSA©: A novel Python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research (Santamaría-Vázquez et al., 2023).

Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Diego Marcos-Martínez, Víctor Rodríguez-González Sergio Pérez-Velasco, Selene Moreno-Calderón, and Roberto Hornero. *Computer Methods and Programs in Biomedicine*, vol. 230, 2023. Impact factor in 2021: 7.027, Q1 in "Computer Science, Interdisciplinary Applications", "Computer Science, Theory & Methods", "Biomedical Engineering" and "Medical Informatics" (JCR-WOS).

Abstract: Background and objective. Neurotechnologies have great potential to transform our society in ways that are yet to be uncovered. The rate of development in this field has increased significantly in recent years, but there are still barriers that need to be overcome before bringing neurotechnologies to the general public. One of these barriers is the difficulty of performing experiments that require complex software, such as brain-computer interfaces (BCI) or cognitive neuroscience experiments. Current platforms have limitations in terms of functionality and flexibility to meet the needs of researchers, who often need to implement new experimentation settings. This work was aimed to propose a novel software ecosystem, called MEDUSA©, to overcome these limitations. *Methods.* We followed strict development practices to optimize MEDUSA© for research in BCI and cognitive neuroscience, making special emphasis in the modularity, flexibility and scalability of our solution. Moreover, it was implemented in Python, an open-source programming language that reduces the development cost by taking advantage from its high-level syntax and large number of community packages. *Results.* MEDUSA© provides a complete suite of signal processing functions, including several deep learning architectures, and several ready-to-use BCI and neuroscience experiments, making it one of the most complete solutions nowadays. We also put special effort in providing tools to facilitate the development of custom experiments, which can be easily shared with the community through an app market available in our website to promote reproducibility. *Conclusions.* MEDUSA© is a novel software ecosystem for modern BCI and neurotechnology experimentation that provides state-of-the-art tools and encourages the participation of the community to make a difference for the progress of these fields. Visit the official website at https://www.medusabci.com/ to know more about this project.

1.2 Dissertation Context

This dissertation is focused on the improvement of several aspects of BCIs to increase their applicability outside the laboratory. Therefore, it falls within the area of biomedical engineering, and more specifically, it is related to the fields of biomedical signal processing, neural engineering and assistive technologies. In the following paragraphs, an overview of these disciplines is presented in order to provide a common framework for the study.

The field of biomedical signal processing studies how to measure and analyze physiological signals (Bronzino and Peterson, 2014). Our organism is composed by a great number of systems (e.g., nervous, cardiovascular, gastrointestinal, res-

piratory) that work together to maintain our body function. Through different mechanisms, we can acquire and analyze physiological signals that measure some feature of these systems over time in order to gain knowledge or perform diagnostics (Bronzino and Peterson, 2014). Some of the greatest discoveries in this discipline occurred as early as the early 20th century, the most notable being the invention of the electrocardiogram (ECG) in 1903 by William Eindhoven (Bronzino and Peterson, 2014). Another of the great milestones in this field was the EEG, discovered in 1929 by Hans Berger (Berger, 1929). During the last decades, this field experienced a rapid expansion, and nowadays there are tens of physiological signals that are widely used for different purposes. These signals are usually classified based on their origin, being the most representative: bio-electrical, bio-impedance, bio-acoustic, bio-mechanic and bio-chemical (Bronzino and Peterson, 2014). In this work, we applied biomedical signal processing to analyze the EEG, which measures the electrical activity of the brain. Thus this field played a crucial role in this work.

Another field that is related to this investigation is neural engineering, also known as neuroengineering. This emerging discipline encompasses any technology that combines principles from neuroscience and engineering to understand, design, diagnose, replace, repair, or enhance biological or artificial neural systems (He, 2020). This general definition encompasses a very wide range of fields, including neuroimaging, neuromodulation, neuromechanics, neural prostheses, neuromorphic circuits, neural tissue engineering or computational neuroscience, among others (Durand, 2006). BCIs are also included in this definition, using neural engineering techniques to measure brain activity and identify specific patterns or waveforms that can be used to decode the user's intentions and control an application (He, 2020). These waveforms are called control signals, and the study and optimization of the cognitive tasks or the stimulation paradigms that trigger them is one of the main research areas in the BCI field. Concretely, this doctoral dissertation focused on visual ERPs, which are the response to visual stimuli reflected in the EEG. In this regard, ERPs are a reliable and widely studied control signal that can be used for assistive applications.

Finally, assistive technologies are products, equipment, and systems that are used to increase, maintain, or improve the functional capabilities of persons with disabilities. Some examples of assistive technologies are wheelchairs, special-purpose computers, prosthetics, screen readers, power lifts, eye trackers, etc. BCIs were also conceived as assistive technologies, and despite that the range of uses has increased in recent years, this application remains as one of its main purposes

(Wolpaw et al., 2002). In fact, BCIs have the potential to solve many of the communication and dependency problems of severely disable people once the technical limitations of these systems are resolved. This dissertation focused on this application to design, develop and test the contributions that have been proposed in the four papers of the compendium, and thus, it represents the higher level framework of the investigation.

1.3 Brain–computer interfaces

Ever since the development of the EEG in 1929 by Hans Berger, scientists have speculated about the use of BCIs to decode neural activity in real time to provide new forms of interaction with the outside world. A BCI capable of interpreting thoughts, desires, or intentions could drastically change the way we interact with our environment, unlocking an unimaginable world of possibilities —and risks— in multiple fields. Although we are still far from this point, the field has rapidly evolved in the last decades to build increasingly complex and accurate BCI systems.

Formally, BCIs are systems that provide an alternative communication pathway between the user and the environment by decoding brain activity in real time to replace, restore, enhance, supplement or improve the human natural outputs, i.e., motor responses (Wolpaw and Wolpaw, 2012). In general, all BCIs have a common workflow with three stages: (1) recording of neural activity with neuroimaging techniques, such as the EEG; (2) decoding of recorded data using signal processing to detect the user's intentions; and (3) translation of these intentions into application commands and execution of these commands. This general workflow is depicted in Figure 1.2. As can be seen, BCI systems form a closed loop. The user's intentions, encoded in their brain activity, drive responses that have a tangible impact on the environment through the BCI system, which leads to further interactions. Therefore, this workflow follows the rules of operant conditioning, a crucial feature to achieve natural interactions (Wolpaw and Wolpaw, 2012).

The next subsections provide an overview of the most important topics to design and develop modern BCIs, defining a common framework from which to approach the rest of the dissertation.

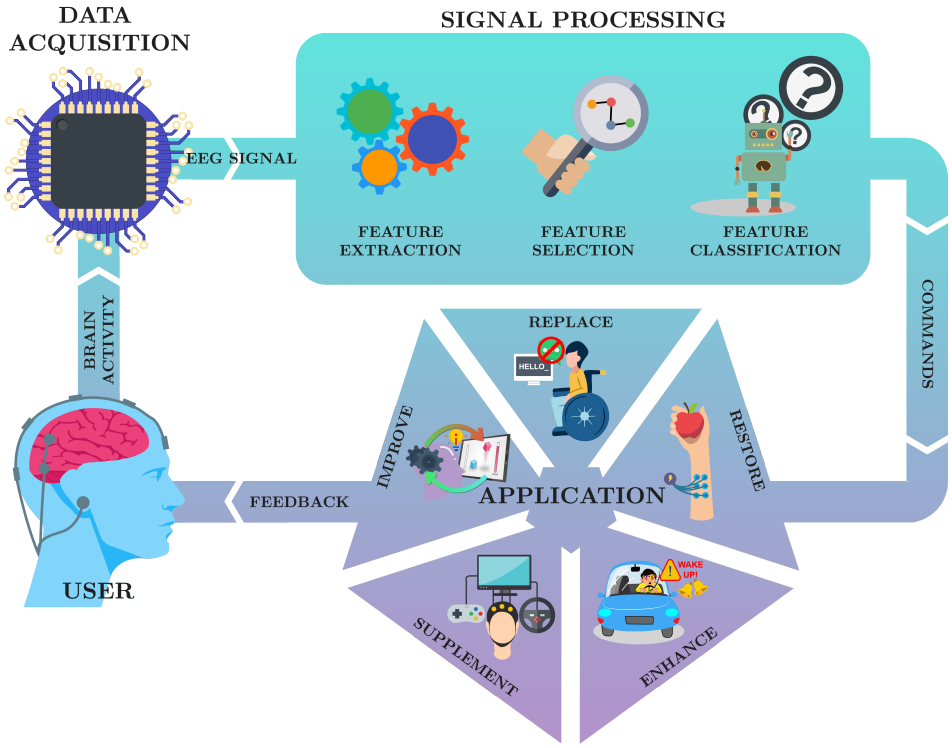

Figure 1.2: Schematic representation of a EEG-based BCI system. Firstly, the signal is acquired with electrodes over the scalp. Then, the signal processing stage decodes the user's intentions and translated them into commands. Finally, the commands are sent to the final application, which provides feedback to the user. BCI applications can replace, restore, enhance, supplement and/or improve the natural body outputs. Adapted with permission from Martínez-Cagigal (2020)

1.3.1 Recording neural activity

The central nervous system is the most complex part of multi-cellular organisms. It can be seen as a control center that receives information from the outside world and acts accordingly. Neurons are the functional units of the nervous system, thus representing a central part in its study (Jessen, 2004). Figure 1.3 depicts a typical neuron, which consists of a cell body, or soma, dendrites and the axon. These specialized cells are specifically designed to receive and generate electrical signals. Between the intra-cellular and extra-cellular spaces, there is a gradient of energy, called resting membrane potential, created by the different concentration of positive and negative ions on both sides of the cellular membrane (Raghavan et al., 2019). The cell can control the resting membrane potential by changing

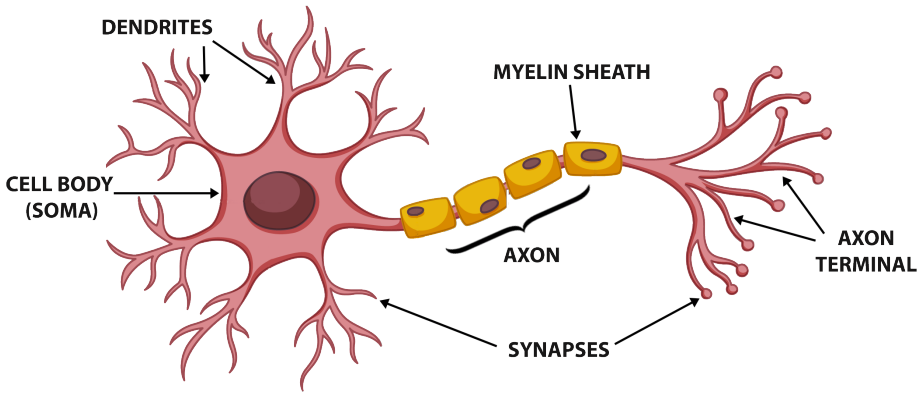

Figure 1.3: Schematic representation of a neuron. Dendrites receive electrical impulses from other neurons. The cell body, or soma, is the core section of the neuron, maintaining its structure and providing energy to drive function. The axon transmits electrical impulses to the axon terminals, and can be coated in myelin, a protein that improves the transmission. Dendrites and axon terminals are connected by synapses. Adapted from He (2020).

the concentration of the ions across the membrane, predominantly sodium (Na^+), potassium (K^+), and chloride (Cl^-) (Raghavan et al., 2019). Since the cell lipid bilayer membrane is not permeable to ions, this change is enabled by specialized structures called ion channels and ion pumps (Raghavan et al., 2019). Ion channels allow the passive movement of ions down their electrochemical gradient. On the other hand, ion pumps transport ions against their natural gradient at the expense of energy (Raghavan et al., 2019). In normal conditions, the resting membrane potential has a negative value between -40 and -80 mV relative to the extra-cellular medium. However, under certain conditions of excitation, neurons trigger regenerative events that revert the resting potential to values close to +50 mV, returning to the baseline point after a few milliseconds. These events, called action potentials, are propagated through the neural network, being central for multiple functions, including information processing and transmission, muscle contraction or hormone release (Raghavan et al., 2019). A neuron receives actions potentials from other neurons through the dendrites and cell body, and transmits new stimulus through the axon (Raghavan et al., 2019). The interaction between axons and dendrites is governed by complex biochemical interactions in the synapses. Whether or not an action potential is generated in the cell body of a neuron depends on the input signals, forming a very complex communication system (He, 2020). However, it should be noted that there are still many open

questions regarding the details of the structure and functioning of neurons. In fact, it is estimated that the human brain roughly 100 billion neurons, with more than 40 different types described to date, each one of them with different characteristics that are not always fully understood (He, 2020).

From a macroscopic point of view, neurons arrange in networks that, together, form the nervous system. Particularly, the central nervous system (CNS) is composed by the spinal cord and the brain, being the latter the organ responsible for the centralized control of the majority of body functions. Anatomically, the brain can be divided in three parts, as shown Figure 1.4: brainstem, cerebellum, and cerebrum (Jacobson, 2008). The brainstem connects the inferior parts of the brain with spinal cord, having critical functions such as the transmission of sensory and motor information, cardiac and respiratory regulation, or sleep cycle management. The cerebellum is related to accurate unconscious motor control (timing, coordination and body posture), and primary cognitive functions and emotions, such as attention, emotional control, and fear/pleasure responses. Finally, the cerebrum, which the largest part of the human brain, contains the cerebral cortex and subcortical structures like the hippocampus, basal ganglia and olfactory bulb. This part of the brain controls the majority of high-level functions, being the cerebral cortex the most important structure for the purposes of this dissertation.

The cerebral cortex, whose main lobes are shown in Figure 1.4, can be divided in different functional areas, being the most relevant for this work (Jacobson, 2008): (1) the primary motor cortex, located in the frontal lobe, plays an important role in motor planning and execution through efferent pathways; (2) the somatosensory cortex, located in the parietal lobe, receives sensory information (e.g., temperature, pain, touch or proprioception) through afferent pathways; (3) the visual cortex, located in the occipital lobe, processes the visual information received through the optic nerves; (4) the auditory cortex, located in the superior and medial areas of the temporal lobes, processes sounds, being also involved in language; and (5) the prefrontal cortex, located in the frontal part of the brain, is fundamentally involved in high-level cognitive processes such as working memory, planning motor execution, decision making, personality and social behaviour, language, etc.

When the brain performs a cognitive task, the neurons involved in the process increase their activity. Typically, complex cognitive processes activate multiple local networks that are sensitive to the current task. The neurons that form these networks synchronize their activity and increase the firing rate of action potentials to process and transmit relevant information. At a higher level, these local networks also synchronize through long distance connections with other parts

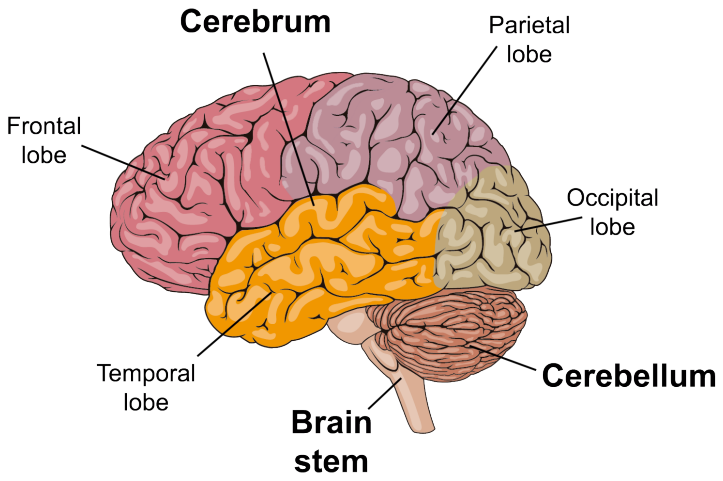

Figure 1.4: Schematic representation of the brain, divided in its three anatomical parts: brainstem, cerebellum and cerebrum. The lobes of the cerebral cortex are also displayed. Adapted from Jacobson (2008).

of the brain, forming a larger network that leaves a recognizable activity footprint. This allows to study which parts of the brain are involved in a certain cognitive task (Jacobson, 2008).

The goal of a BCI is to decode this activity in real time to detect the user's intentions. Therefore, in order to implement a BCI, the first step is the acquisition of data reflecting the underlying neural activity of the user's brain, and especially, the cortex. Unfortunately, this measurement is a difficult challenge. First, the physical barriers that protect the brain (e.g., scalp, skull, dura, arachnoid) make it difficult to measure neural activity directly on the source, which would maximize precision (Wolpaw and Wolpaw, 2012). In fact, this is only possible with highly invasive and complex techniques that, for safety and ethical reasons, can only be used as last resource in life threatening situations. Another reason is the exceptional complexity of the brain. The accurate decoding of high-level mental processes would be required to measure the activity of billions of neurons at the same time throughout the brain, but current technology lack the required capacity and precision. However, the development of neuroimaging in the last decades brought up several techniques that opened a window into the brain, enabling crucial advancements for a better understanding of this organ (Wolpaw and Wolpaw, 2012).

Neural activity involves electromagnetic, chemical and metabolic processes. Depending on which one of these processes is measured and how, the neuroimaging method will have different characteristics and purposes. In this regard, practical BCI technology should be precise, non-invasive, cheap, portable, easy to use, and comfortable to reach a wide acceptance by both end-users and clinicians.

The precision will depend on three important aspects of the neuroimaging technique: the spatial resolution, temporal resolution and coverage. Spatial resolution refers to the minimum number of neurons whose activity can be recorded with a neuroimaging method. For instance, a method with perfect spatial resolution would allow to record the activity of single neurons. On the other hand, temporal resolution accounts for the fastest event that be detected. Ideally, a method with perfect temporal resolution would register any change in the neuron's activity, including rapid variations on the membrane potential or ions concentration. Finally, the coverage is the volume of brain covered by the recording. A neuroimaging technique with high spatial and temporal resolution, but that only covers a small part of the brain (e.g., 1 mm^3) cannot be used to decode high-level cognitive processes, since the processing is distributed across many parts of the brain.

The next crucial characteristic that affects the election of a neuroimaging technique to build a BCI is the degree of invasiveness. In this regard, invasive techniques should only be considered for life-threatening situations and extreme cases (e.g., locked-in patients). However, for applications such as the majority of assistive systems (leaving aside extreme cases), neurorehabilitation or entertainment, invasive systems are not a viable option, especially if they involve neurosurgery.

Finally, aspects such as the price and the usability of the BCI should also be considered, especially for assistive systems and entertainment applications. In the following subsections we review the most widely used neuroimaging methods nowadays, analyzing these aspects. According to their underlying principle, current techniques can be divided in two categories:

Metabolic

Neural activation due to cognitive activity increases the firing rate of action potentials and other processes that require energy. This energy is provided by oxygen and glucose, the basic fuels of metabolism. These components are supplied to the neurons by circulatory system through capillaries. When neurons require more energy, the blood flow to that region is increased to meet the demands. This change in the hemodynamic response of the brain is locally regulated, thus providing a marker for neuronal activity. Currently, the most important techniques that use

metabolic processes to record the brain's activity are (Sitaram et al., 2012):

1. **Functional transcranial Doppler (fTCD).** This technique uses pulse-wave doppler technology to measure cerebral perfusion changes due to neural activation. The main advantages are its simplicity and portable equipment. However, its low spatial resolution only allows to measure changes in major arteries between different hemispheres of the brain. This drawback prevents the use of fTCD in BCI, but since its development it has substantially contributed to our understanding of the hemispheric brain organization (Baumgartner, 2006; Sitaram et al., 2012).

2. **Positron emission tomography (PET).** This technology traces radioactive substances to visualize and measure biochemical changes. It requires the injection of radioactive compounds, called radiotracers, which bind to different chemicals. PET traces these substances to detect changes or anomalies in the distribution of the chemical throughout the body. In neuroscience, PET combined with different radiotracers (e.g., oxygen-15, F-FDG) allows to detect changes in metabolism. Despite the usefulness of this method, which is widely used to detect tumors or infections in the daily clinical practice, it is of no interest for BCIs due to its cost, low temporal resolution, non-portable equipment, and slight invasiveness (Sitaram et al., 2012).

3. **Functional near-infrared spectroscopy (fNIRS).** This method detects changes in the oxygenation levels of the blood using infrared light of different wavelengths. These changes, known as the blood-oxygen-level-dependent (BOLD) response, are measured by tracking different oxy- and deoxy-hemoglobin concentrations from the surface of the cortex (Sitaram et al., 2012). fNIRS represents a feasible candidate for certain BCI systems due to its coverage, non-invasiveness, and portability (Naseer and Hong, 2015). Nevertheless, its low temporal resolution, limited by the slow metabolic response, and low spatial resolution, make this technique unsuitable for BCIs that require multiple control choices and high information transfer rates (ITR) (Naseer and Hong, 2015).

4. **Functional magnetic resonance imaging (fMRI).** As the fNIRS, this technique also measures the BOLD response. In this case, the measurement is based on the distinctive response of oxy- and deoxy-hemoglobin to high-power magnetic fields, allowing to substantially increase the spatial resolution with respect to fNIRS (Sitaram et al., 2012). On the other hand, it has

low temporal resolution and it requires expensive, non-portable equipment. For these reasons, although fMRI can be considered the most suitable technique based the measurement of metabolic processes to design certain types of BCIs, it is impractical for most situations and the range of applications is mostly limited to neurorehabilitation (Yuan et al., 2021).

In general, methods based on the detection of metabolic processes are non-invasive (except PET), achieve good spatial resolutions that range between less than 1 mm^3 to 5 mm^3, and allow a broad brain coverage. On the other hand, the hemodynamic response is slow, taking at least 2s for a blood vessel to react to increased glucose and energy demands. Therefore, the temporal resolution is limited, hindering the detection of brief, sparse events. Despite this drawback, blood-flow imaging has been instrumental in many advancements in neuroscience, being the most widely used techniques in this field.

Electromagnetic

The number of studies that used metabolic methods, especially fMRI, to map cognitive functions into the underlying neuroanatomy of the brain grew exponentially in the last decades. Despite the undeniable advances that this approach has provided in linking cognitive functions to different areas of the brain, we now know that it can only show a part of the whole picture. This, perhaps, overemphasis on the spatial mapping approach was probably inherited from the pre-neuroimaging era, when brain function was primarily understood through localized brain lesions. However, recent advances in neuroimaging techniques with high temporal resolution showed that a precise comprehension of the spatio-temporal dynamics of neural activity throughout the brain can fill the gaps in our current understanding.

These techniques are based on the measurement of the electrical and magnetic fields generated by neurons, which are assumed to be the primary form of communication in our nervous system. The current methods are (Srinivasan, 2012):

1. **Local fields potentials (LPF)**. This signal is the electric potential recorded in the extracellular space of brain tissue. Typically, it is recorded with micro-electrodes, such as metal, silicon and glass micropipettes, which are implanted in the brain, reflecting the activity within volumes of 10^{-3} mm^3 to 1 mm^3.

2. **Electrocorticography (ECoG)**. This method uses electrodes placed over the surface of the cortex to measure the activity at the mesoscopic scale,

targeting a volume of brain tissue that ranges between 1 mm^3 and 20 mm^3.

3. **Electroencephalography (EEG)**. This technique uses electrodes placed over the scalp to record the electrical activity of the brain. It is estimated that a single EEG electrode measures the activity within tissue volumes between 10^3 mm^3 and 10^4 mm^3, reflecting the averaged electrical potentials of a hundred million to a billion of neurons (Srinivasan, 2012).

4. **Magnetoencephalography (MEG)**. This signal reflects the magnetic activity of the neurons at the macroscopic scale. Due to the small amplitude of the magnetic field relative to unavoidable ambient magnetic variations, its measurement requires the use of a superconducting quantum interference device magnetometer at a very low temperature placed in a magnetically shielded chamber (Srinivasan, 2012). Noteworthy, despite that EEG and MEG reflect neural activity at the same scale, electric and magnetic fields generated by the neurons are partly independent due to the low frequency of brain signals and specific source characteristics. This implies that MEG and EEG are sensitive to the activity of different neurons, providing complementary information (Srinivasan, 2012).

The previous techniques are based on the same underlying physiological principles, being the main differences between them the way to measure the electromagnetic fields and the location and characteristics of the electrodes. All of them present excellent temporal resolution due to the propagation speed of electromagnetic fields. Regarding the spatial resolution, LPF achieves the best spatial resolution, followed by ECoG, MEG and EEG, which has the lowest one. Moreover, LPF, ECoG and EEG use portable equipment with a relatively low cost compared to metabolic neuroimaging and MEG. With respect to the degree invasiveness, EEG and MEG are the only ones that are non-invasive and completely safe, whereas the other techniques present a high risk for the subject. This allows to place as many electrodes as needed over the scalp, providing a total coverage of the cortex's surface. In fact, modern high-density EEG and MEG equipment records up to 256 channels at different positions at the same time.

Taking into account all the aspects that should be considered for designing BCI systems, the EEG is the best technique in most scenarios. The reasons are its excellent temporal resolution, total coverage of the cortex's surface, non-invasiveness, price, and usability, being the weakest point its low spatial resolution. Other techniques that can be used in certain scenarios are the ECoG, which can be used

in extreme cases with severely disabled subjects, and fNIRS/fMRI in some neurorehabilitation applications. However, the EEG is, by far, the most extended neuroimaging technique in the BCI field and it is the one that was used in this dissertation (Wolpaw and Wolpaw, 2012).

1.3.2 EEG

EEG records electric potentials using electrodes placed over the user's scalp. The origin of these potentials is the neural activity of the brain, but the underlying generation mechanisms are not yet fully understood (He, 2020). The most accepted hypothesis is that the EEG reflects averaged oscillations provoked by postsynaptic activity of large populations of pyramidal neurons from the cortex (He, 2020). On the other hand, the contribution of action potentials is questioned for two reasons: (1) their short generation time (i.e., $\sim$ 1ms), which makes synchronization between neuronal assemblies difficult, thus canceling the resulting electric field measured from a macroscale point of view; and (2) the strong frequency-filtering properties of the tissues separating sources from the electrodes, which remove frequencies above 100 Hz (He, 2020). Additionally, the volume conduction problem has to be taken into account. This phenomenon accounts for the spread of the electric field when it crosses through different layers of tissue, decreasing the spatial resolution of the EEG. Therefore, this signal only reflects a small, but important part of the underlying neural activity. Despite the fact that EEG showed consistent and strong correlation with brain states, these limitations hide the physiological origin of the recorded oscillations. However, the increasingly complex methods to analyze EEG, which have been developed in the last decades, are improving our understanding of this signal, increasing its presence in applied clinical and neurotechnology settings.

EEG is composed by spontaneous and evoked activity. Spontaneous activity is present in the absence of explicit inputs (e.g., sensory stimuli) or outputs (e.g., motor responses) (Musso et al., 2010). This activity is attributed to both conscious and unconscious processing. There is increasing evidence to support that these oscillations are organized in a coherent default-mode network with a specific spatio-temporal structure that helps to maintain homeostasis, which is necessary for proper brain function (Musso et al., 2010). On the other hand, evoked activity is triggered by specific events, such as a sudden sound or the movement of an arm. Thus, it is time-locked to events, having a specific temporal onset, and leaves characteristic footprints on the EEG that can be replicated under the same experimental conditions. For this reason, EEG patterns associated with evoked ac-

tivity can be extracted by averaging different EEG features (e.g., amplitude, band power, etc.) for multiple trials of the same experiment to cancel the influence of the spontaneous activity, which is superimposed and follows a random distribution (Wolpaw and Wolpaw, 2012). The result of this synchronized average is an ERP, which is composed by different components that can be classified in three main categories (Luck, 2014): (1) exogenous components provoked by primary processing of sensory stimuli; (2) endogenous components triggered by mental tasks and high-level cognitive processes; (3) motor components that accompany the preparation and execution of movements. It should be noted that different events leave different footprints with specific spatio-temporal structure.

Due to the complexity of EEG, it is difficult to extract meaningful information from this signal simply by visual inspection in the temporal domain. EEG reflects the synchronized activity of large populations of neurons that engage in different cognitive processes at the same time, often behaving as an oscillatory phenomenon with a specific spatio-temporal structure. For this reason, it is often represented in the frequency domain, which allows to analyze the underlying rhythmic activity. Numerous studies showed that certain brain processes or mental states are associated with an increased/decreased activity in different frequency bands. Typically, the EEG is divided into 5 bands, according to their characteristics and frequency ranges: delta ($\boldsymbol{\delta}$, 1–4 Hz), theta ($\boldsymbol{\theta}$, 4–8 Hz), alpha ($\boldsymbol{\alpha}$, 8–13 Hz), beta ($\boldsymbol{\beta}$, 13–30 Hz), and gamma ($\boldsymbol{\gamma}$, 30–100 Hz) (Tantum, 2021).

An EEG recording typically needs electrodes to capture the electrical potentials and a biomedical amplifier to amplify, filter and convert the EEG signal to the digital domain. The minimum setup requires at least three electrodes: ground, reference, and signal channel. The ground electrode is connected to the amplifier ground and isolated from the power supply to prevent drifts and optimize common-mode rejection. Then, the EEG signal is obtained as the difference between the channel and the reference, which is typically placed far the brain sources, such as the ear, mastoid, or neck (Srinivasan, 2012). Clinical and BCI experiments often require multiple EEG channels. In order to achieve comparable, consistent and reproducible EEG recordings, the American Clinical Neurophysiology Society proposed several standards to define the positions of the electrodes, being the most widely used are the International Systems 10-20 and 10-10 (Acharya et al., 2016). These systems are based on the average distance between the inion and the nasion, which are used as reference points. The Figure 1.5 depicts the International System 10-10.

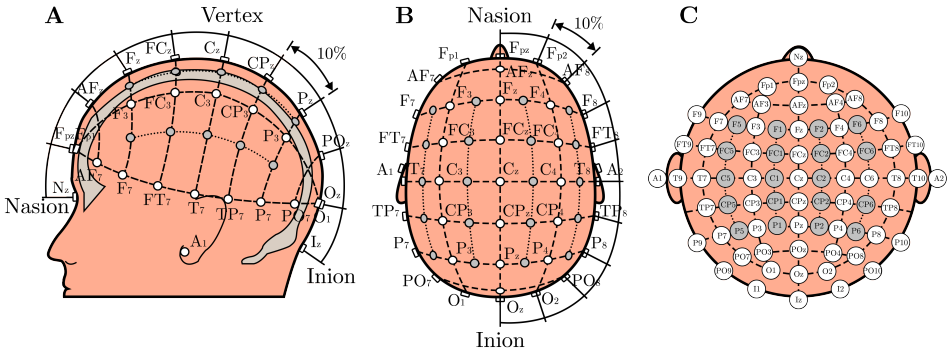

Figure 1.5: The International System 10-10 illustrated in schematic views showing (A) a lateral view, (B) a top view, and (C) a scalp projection. This standard uses proportional distances of 10% between the nasion and the inion for positioning the electrodes on the scalp (Martínez-Cagigal, 2020).

1.3.3 BCI control signals

Decoding neural activity from EEG signals is not straightforward due to the low spatial resolution and high complexity of this signal. To achieve this goal, EEG must first encode the user's intention that is being conveyed. Then, different signal processing techniques must be applied to isolate this information and translate it into an application command.

In order to encode user's intentions in the EEG, BCIs use certain types of evoked activity. During the last decades, researchers studied different paradigms to optimize the use of evoked activity for BCI control, using neural engineering techniques to trigger adequate responses for each application. These responses are known as control signals, being the most important nowadays:

1. **Sensorymotor rhythms (SMR)**. SMR are components recorded in the sensorimotor cortex in alpha, beta, and gamma bands. SMR activity is provoked by the conscious preparation of a motor intent and is characterized by: (1) a contralateral power decrease, known as event-related desynchronization (ERD); and (2) an ipsilateral power increase, known as event-related synchronization (ERS) (Pfurtscheller and McFarland, 2012). ERD/ERS events are bound to the area of the cortex that controls the muscles involved in the movement. An interesting feature for BCI control is that, since SMRs are modulated by motor planning and attention mechanisms, instead of actual movement execution, executed and imagined movements can trigger similar ERD/ERS responses. Therefore, this control signal can be used by different

types of users, including severely disabled, to control a BCI system. However, in order to achieve a successful control through motor imagery exercises, users usually need to train for several hours. Another handicap of the SMRs is that the number of classes that can be discriminated is limited due to the low spatial resolution of the EEG. Therefore, the number of different commands that can be implemented in a SMR-based BCI is also limited. In fact, current SMR-based BCIs rarely implement more than 4 commands to maintain a suitable accuracy.

2. **Steady-state visual evoked potentials (SSVEP)**. SSVEPs are the result of rapid and repetitive visual stimulation at fixed rates (Luck, 2014). In this paradigm, the resulting evoked responses are overlapped producing a constant oscillation at the stimulation frequency. Since these evoked responses are the result of low-level sensory processing, no mental tasks are involved. Therefore, SSVEPs are a considered an exogenous control signal that can be triggered in the user without training. If the stimulation is maintained enough time, SSVEPs can be identified with a simple frequency analysis. The classical setup of SSVEP-based BCIs is a matrix with cells that produce visual stimuli at different frequencies. When the user pays attention to a specific cell, the SSVEP is detected in the EEG and the corresponding application command is executed. This setup allows to implement a high number of commands (typically between 12 and 64), but it is still limited by the refreshing rate of current LCD screens (Luck, 2014).

3. **Code-modulated visual evoked potentials (c-VEP)**. c-VEPs are elicited by flashing stimulation sequences that follow pseudo-random noise codes, generating evoked neural activity correlated with the sequence the user is paying attention to (Martínez-Cagigal et al., 2021). In this paradigm, several cells are arranged in a matrix, each one of them with a different stimulation sequence associated to a command. When the user wants to perform a selection, the BCI system detects which sequence is more correlated to the EEG activity and executes the corresponding command (Martínez-Cagigal et al., 2021). In recent years, this paradigm achieved the highest accuracy and selection speed ever reported in the BCI literature, being one of the most promising alternatives for communication and control BCIs. However, c-VEP-based BCIs require complex time synchronization and the number of commands, typically between 4 and 16, is still limited by the availability of uncorrelated psudo-random codes and the refreshing rate of LCD screens

(Martínez-Cagigal et al., 2021).

4. **P300 evoked potentials**. In this paradigm, the user is presented with different cells, each one associated to a command, that are highlighted sequentially in random order. The task of the user is to focus on the target cell while ignoring the other stimuli that are presented. In this setup, target events elicit visual ERPs with the P300 potential in the EEG, whereas the other stimuli triggers visual ERPs without this component, as can be seen in Figure 1.6 (Santamaría-Vázquez et al., 2019a). The P300 component is a positive potential elicited by the recognition of a rare stimulus (target) within a series of frequent stimuli (non-target), appearing around 300 ms after the stimulus onset. In order to decode the command, the BCI system detects the P300 in the EEG using signal processing techniques. This waveform is considered an endogenous component, since it reflects high-level cognitive tasks to differentiate between the two different types of stimuli, but the user does not require training to elicit it. This control signal has been widely studied to implement systems for communication and control. Despite the fact that P300 evoked potentials generally provide lower accuracy and speed than SSVEPs and c-VEPs, the number of commands is not limited, thus allowing to control complex interfaces such as domotic systems or full keyboards. Furthermore, it does not require complex user interfaces or accurate time synchronization. For these reasons, this control signal is considered the most reliable for applied systems outside the laboratory, probably being the most extended in the BCI field nowadays (Sellers et al., 2012). The systems that use this control signal are known as ERP-based BCIs or P300-based BCIs. Although both terms are often used interchangeably in the literature, in this document we use the term ERP-based BCIs.

1.3.4 EEG processing in BCI

Once the information to discriminate the command has been encoded in the EEG using the previous or other control signals, the system must detect the corresponding neural activity and isolate it from the rest of EEG components by applying signal processing methods. Although the algorithm will depend on the specific BCI paradigm, we can define a general signal processing framework implemented in 4 sequential stages: pre-processing, feature extraction, feature selection and feature classification.

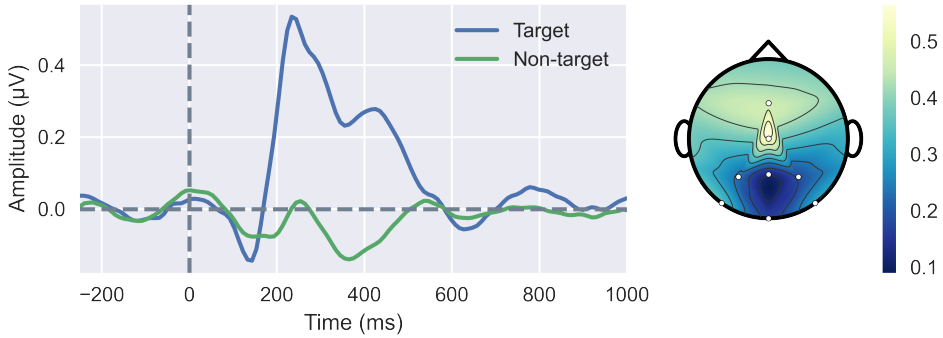

Figure 1.6: The figure illustrates the averaged EEG epochs for both target and non-target stimuli in an ERP-based BCI. The left graph represents the signal in channel Cz, with the stimulus onset placed at t=0. As can be seen, the ERP, with the P300 component, is clearly visible for target stimuli. The graph on the right shows the amplitude values recorded in eight different channels at the peak of the ERP.

The goal of the pre-processing stage is to remove non-target components and noisy artifacts from the EEG, increasing the signal-to-noise ratio (SNR) of the control signal (Lotte et al., 2018). The most typical methods include: (1) frequency filters such as infinite impulse response (IIR) filters and finite impulse response (FIR) filters to remove frequencies outside the interest band; (2) spatial filters to increase the spatial resolution, such as common average reference (CAR), Laplacian Filters or Common Spatial Patterns; and (3) source separation to remove certain signal components, such as Independent Component Analysis (ICA) (Lotte et al., 2018).

In the feature extraction stage, the algorithm extracts meaningful features to detect the control signal in the EEG, facilitating the learning and generalization steps in the classification stage (Lotte et al., 2018). Depending on the control signal, different features can be used, being the most relevant: (1) temporal features such as amplitude distribution indices or moving averages; (2) spectral features based on Fourier analysis, such as band powers or spectral entropy; (3) non-linear metrics such as central tendency measure (CTM), Lempel-Ziv complexity (LZC) or Sample entropy (SampEn); (4) connectivity metrics, including amplitude and phase parameters such as phase locking value (PLC), amplitude envelope correlation (AEC); and (5) graph theory metrics. A common approach is to calculate different types of features and use all of them in the following stages (Lotte et al., 2018).

After the feature extraction stage, the information contained in the EEG is syn-

thesized in a set of variables (Lotte et al., 2018). However, some of these features may provide irrelevant or redundant information for the control signal detection problem, which may reduce the performance of the algorithm and increase overfitting. Therefore, an optional feature selection stage can be applied to select the most relevant features. There are two types of algorithms that can be used. Filter approaches are independent from the classification stage, selecting the most relevant features based on correlation analyses. On the other hand, wrapper methods use predictions with a given model to decide which features are more relevant for the task, calculating the optimal set for a specific classification method (Guyon and Elisseeff, 2003).

The feature classification stage automatically detects the EEG patterns corresponding to the control signal depending on the values of the input features. The classification method calculates the optimal decision boundaries in the feature space to separate the classes based on different criteria. For EEG classification tasks, the most successful methods are based on supervised learning, which require a training set to calculate the optimal decision boundaries (Lotte et al., 2018). Some widely used classifiers for EEG classifications tasks are linear regression, linear discriminant analysis (LDA), support vector machines (SVM), artificial neural networks (ANN) and Riemannian-based classifiers (Lotte et al., 2018). Depending on the paradigm, an additional step might be necessary to translate the classifier output to application commands, as in ERP-based BCIs, where the detected P300 must be mapped to their corresponding command.

1.3.5 BCI Applications

A BCI system works as a general interface between the user's brain and an external device. Therefore, this technology can be potentially used for unlimited applications. Nevertheless, due to its current high complexity and low performance, its use is restricted to specific fields. Nowadays, the most important applications of BCIs on the biomedical field are assistive systems and neurorehabilitation.

In the 1980s, BCIs were conceptualized as an aid system for people with communication and motor control problems (Farwell and Donchin, 1988). Nowadays, increasing the quality of life and independence of severely disabled people is still one of the main goals of the EEG-based BCI technology (Wolpaw and Wolpaw, 2012). In the last decades, many different kinds of assistive systems controlled through BCI appeared: web browsers (Martínez-Cagigal et al., 2017), wheelchairs (Cruz et al., 2021), exoskeletons and robotic limbs (Tariq et al., 2018), social media

applications (Martínez-Cagigal et al., 2019a) or alternative communication devices (Vansteensel and Jarosiewicz, 2020).

BCIs can also be used to rehabilitate or enhance brain functions through repetitive endogenous stimulation of brain activity (Enriquez-Geppert et al., 2017). Using neurofeedback (NF) and operant conditioning, BCIs allow to self-modulate neural activity and promote brain plasticity using real-time feedback of neural activation parameters (Enriquez-Geppert et al., 2017). NF has emerged in recent years as a promising methodology to correct pathological states and behaviours, such as attention deficit hyperactivity disorder, anxiety or depression, and even to improve recovery expectancies after suffering brain damage provoked by stroke or traumatic brain injury. Furthermore, NF techniques have been proposed as a cognitive enhancement tool for healthy subjects (Enriquez-Geppert et al., 2017).

1.4 State of the art

The contributions of this dissertation cover a wide range of topics in the BCI field, including: (1) improvement of asynchronous control of ERP-based BCIs; (2) application of novel pattern recognition techniques based on deep learning to improve BCI performance; and (3) the development of a novel platform to accelerate BCI and cognitive neuroscience research. The following subsections provide a comprehensive state of the art review of these topics.

1.4.1 Asynchronous control of ERP-based BCIs

As mentioned in previous sections, BCIs use evoked activity to decode the user's intentions from the EEG. However, in a single EEG observation, control signals (e.g., SMR, ERP) are, in most cases, barely discernible from the spontaneous components due to the low SNR of this signal. For this reason, the command selection time is usually pre-allocated. This approach allows to synchronize neural events in order to increase the performance of the signal processing algorithm, simplifying the classification task. The main implication is that the user can only interact with the system during specific, fixed-duration time windows. Furthermore, to maximize the accuracy, the system assumes that the user always wants to make a selection in the given time slot. Thus, it always decides and executes a command even if the user does not want to interact with the BCI.

This synchronous behaviour is not a limitation in neurorehabilitation applications, where the goal is to facilitate endogenous stimulation of certain neural

activity patterns. But for practical assistive applications, this poses a great drawback. In this case, users must be able to shift their attention away from the BCI to attend different tasks on demand (Santamaría-Vázquez et al., 2019a). If the system randomly selects commands when the user is not controlling it, it can potentially create undesired or even dangerous situations, as in an application for wheelchair control. In these cases, the system needs to implement an additional control state detection layer to provide an asynchronous control of the application by dynamically detecting whether the user is willingly controlling the BCI (i.e., control state), or is attending to other task (i.e., non-control state). As one of the most widely used paradigms for assistive applications, asynchronous ERP-based BCIs based on the P300 potential are of great interest.

Several works studied different strategies to provide an asynchronous control of ERP-based BCIs. The approaches can be divided into two categories: (1) thresholds dependent on the output scores of the feature classification stage (Aloise et al., 2011; Aydin et al., 2018; He et al., 2017; Martínez-Cagigal et al., 2017, 2019a; Zhang et al., 2008); or (2) hybrid BCIs that combine different control signals (Li et al., 2013; Panicker et al., 2010; Yu et al., 2017). Nevertheless, these approaches still present limitations. For instance, algorithms that rely on the output scores of the ERP classification stage have high inter-session variability. Small changes in the amplitude or latency of the ERPs could override the threshold and cause a drastic decrease in the accuracy of these methods. In fact, thresholds must be recalibrated before each session with the BCI system. This procedure is time consuming, affects the usability of the system and it is frustrating for users (Martínez-Cagigal et al., 2017; Schettini et al., 2014). Moreover, previous approaches require to record additional data to characterize the EEG when the users are not paying attention to the system in order to calculate the threshold, which increases the duration of the calibration sessions. On the other hand, hybrid BCIs increase the complexity of the system, requiring different stimulation interfaces. In these BCIs, the user has to be able to manage with two or more control signals, making them more demanding. Consequently, hybrid BCIs might be a major challenge for certain users who cannot maintain high levels of concentration (Amiri et al., 2013).

Algorithms independent of the ERP classification stage may help to overcome the previous limitations. In this regard, Pinegger et al. (2015) proposed a groundbreaking method based on the hypothesis that the flashing frequency of the classical P300-based speller should be encoded somehow in the EEG due to low-level processing of non-target visual stimuli. However, the proposed method

only reached an average accuracy of 79.5% in the control state detection task. Moreover, it required to double the duration of the calibration sessions in order to acquire non-control trials. Therefore, there was still room for improvement, and further developments could solve the limitations of the study of Pinegger et al. (2015). In this regard, novel techniques to detect the low-level processing of visual stimuli are promising candidates to provide a real asynchronous control of ERP-based BCIs.

1.4.2 Deep learning in ERP-based BCIs

Another limitation of current BCIs is their performance. For some applications, including assistive systems, the accuracy and command selection speed, which are the two main metrics to measure the performance of a BCI, are far from ideal. In this regard, there are two ways of improving the accuracy and speed of a BCI: (1) improve the control paradigm to elicit stronger responses in the EEG that are easier to detect (i.e., improve SNR); or (2) improve the signal processing stage to increase the sensibility and specificity of the system for a given paradigm. Although both approaches get a lot of attention, the latter is easier to put into practice, taking into account that you can use existing BCI applications and databases to design novel pattern recognition methodologies.

In the case of ERP-based BCIs, classical methods based on feature engineering and machine learning, such as LDA and SVM, have been widely explored (Lotte et al., 2018, 2007). They achieve a reasonable performance, but their robustness is easily compromised by noisy artifacts or inter-subject and inter-session variability (Lotte et al., 2007). Moreover, several studies demonstrated that the performance of these methods is reduced drastically in real environments with severely disabled populations (Martínez-Cagigal et al., 2017, 2019a). Recent developments in tensor classifiers based on Riemannian geometry (RG) and adaptive methodologies have made some improvements, but their accuracy and generalization ability remain limited (Congedo et al., 2017; Lotte et al., 2018).

In recent years, the development of deep learning brought breaking advances in the field of pattern recognition (Craik et al., 2019). Deep neural networks are able to automatically extract complex features from raw data, learning hierarchical representations of the input in different levels of abstraction (Lecun et al., 2015). This ability has revolutionized fields such as computer vision, natural language processing, genomics or drug-discovery (Lecun et al., 2015). In the EEG domain, deep neural networks have been used for ERP, SMR and SSVEP classification,

seizure detection and prediction, sleep stage scoring, mental workload detection, data augmentation or emotion recognition, among others (Craik et al., 2019). However, there are still relatively few studies of deep-learning models for EEG analysis (Roy et al., 2019). Since the application of deep learning to EEG is relatively new, most studies explored simple architectures based on convolutional neural networks (CNN) and recurrent neural networks (RNN) (Craik et al., 2019). Nevertheless, there is still room for improvement, and more complex architectures specifically designed for EEG analysis could enhance the performance of CNNs in this domain.

Several studies proposed deep-learning models for ERP detection in a BCI framework. Table 1.1 presents a summary of previous approaches, including a brief description of their contributions, validation method, testing subjects, and outcomes. It can be observed that CNNs are the most widely used method. Among them, EEGNet, proposed by Lawhern et al. (2018), had great impact for its suitable performance and efficient use of depthwise and separable convolutions to create a robust and lightweight structure. Another approach was the model proposed by Borra et al. (2019), which was a fine-tuned version of EEGNet for ERP detection in subjects with attention-deficit/hyperactivity disorder (ADHD). In this table, we also include our first deep-learning model (not included in this compendium), called CNN-BLSTM, which combines a convolutional layer to extract spatial patterns with 2 recurrent layers based on bidirectional long-short term memory units (BLSTM) to learn temporal patterns (Santamaría-Vázquez et al., 2019b). The use of RNNs for EEG processing is scarce, especially for ERP detection (Craik et al., 2019). The reason for this may be that RNNs have a high computational cost and take longer to train compared to CNNs. However, RNNs are well-suited for processing time series data, making them a viable alternative to CNNs in EEG processing.

It should be noted that previous studies have not tested their deep-learning models on subjects with motor impairments. Furthermore, Cecotti and Gräser (2011) and Liu et al. (2018) only included two and three healthy subjects, respectively. It is well-known that patients often have lower classification accuracy due to factors related to their specific conditions, such as neural damage, attention problems, visual impairment, involuntary tremors, or limited cognitive performance (Martínez-Cagigal et al., 2017, 2019a). Furthermore, these symptoms can vary greatly between individuals, even among those with the same condition, making severely disabled subjects a particularly challenging and heterogeneous group for ERP detection. Therefore, a thorough evaluation of new models in this group is

Table 1.1: Previous deep-learning approaches for ERP-based BCIs

Study	Highlights	Validation	Subjects	Accuracy
Cecotti and Gräser (2011)	First CNN for BCI classification tasks	Intra-subject	2 CS	95%
Manor and Geva (2015)	Spatio-temporal regularization	Cross-subject	15 CS	70%
Liu et al. (2018)	Dropout and Batch normalization	Intra-subject	3 CS	97%
Lawhern et al. (2018)	Depthwise and separable convolutions	Cross/intra-subject	18 CS	90/92%
Santamaría-Vázquez et al. (2019b)	Use of bidirectional LSTM layers	Hybrid	15 ADHD	84%
Borra et al. (2019)	Depthwise and separable convolutions	Intra-subject	15 ADHD	92%

CS: control subjects; CNN: convolutional neural network; ADHD: subjects with attention deficit hyperactivity disorder; MDS: motor disabled subjects; CNN: convolutional neural network; LSTM: long-short term memory; Validation: intra-subject strategies train and test the models with data from the same subject, cross-subject strategies train and test the models with data from different subjects and hybrid approaches combine both techniques; Accuracy: test command decoding accuracy.

necessary to determine their usefulness in assistive BCI applications.

1.4.3 BCI platforms

In order to gain new insights into brain function or test new signal processing algorithms, BCI researchers often need to conduct sophisticated experiments with strict temporal requirements and high-precision synchronization between processes. The implementation of such experiments requires specialized software tools to handle the signal acquisition, data processing, task presentation, and feedback to the user (Wolpaw and Wolpaw, 2012). The design and development of these tools is difficult and time consuming, requiring extensive technical knowledge in several domains. Furthermore, a great flexibility is often required in this research environment to rapidly adjust the applications to particular studies or projects. Unfortunately, this expertise is often out of reach for most neuroscience research groups, who stick to the available software or rely on external entities to develop the required programs for their research, leading to delays and cost increases. This is particularly relevant in closed-loop experiments implemented with BCIs due to their technical complexity (Ciliberti and Kloosterman, 2017; Lopes and Monteiro, 2021).

In this context, software tools and applications specifically designed to facilitate the implementation of neuroscience experiments are of great importance for the progress of brain research and neurotechnology. The impact that this kind of tools could have in particular fields should not be underestimated due to their ability

to speed up experimentation, reduce project costs, and enable the participation of researchers without technical knowledge (Renard et al., 2010). While there is a wide range of signal/image processing toolboxes with state-of-the-art methods, such as EEGLAB (Delorme and Makeig, 2004), Brainstorm (Tadel et al., 2011) or MNE (Gramfort et al., 2013), there are limited options for experimental design and implementation. Most of the current tools have few or none customization options and they can only be applied for specific tasks. Moreover, these programs are often distributed under proprietary terms with no compatibility between biomedical recording equipment from different manufacturers, which limits the opportunity to take advantage from all available resources in a research laboratory.

There have been some attempts to overcome these limitations. Two examples of open-source platforms for neuroscience research with success within the BCI community are BCI2000 (Schalk et al., 2004) and OpenVibe (Renard et al., 2010). Although these platforms have been widely used in BCI studies for years, they have drawbacks that should be considered. Their signal acquisition module is not prepared to handle multiple input signals, which limits their application in some experiments (e.g., collaborative and competitive BCIs). Another important aspect is that their implementation in C++, a complex programming language, is not convenient to keep up with the latest developments in the BCI field. For instance, BCI2000 and OpenVibe do not offer some state-of-the-art BCI paradigms, such as those based on c-VEPs, or signal processing algorithms, such as deep neural networks. Moreover, these platforms lack specific tools to create and share new applications and experiments, which are important functionalities in research environments. As a result, despite the useful help that they provided in the past decades, the community has little opportunity to contribute to their development. These limitations also apply to less-known projects that, in some cases, are barely (or no longer) maintained: BF++ (http://www.brainterface.com) (Bianchi et al., 2003), xBCI (http://xbci.sourceforge.net) (Susila et al., 2010) or Pyff (https://bbci.de/pyff/index.html) (Venthur et al., 2010).

To summarize, we identified a gap in the availability of appropriate BCI software tools to carry out the research we had intended to conduct at the beginning of this work. The lack of modern BCI platforms capable of handling the complexity of these new challenges has hindered progress in the field. The development of BCI platforms with more power and flexibility could facilitate further advancements.

Once the main topics of this dissertation have been introduced in this chapter, the rest of the document is organized as follows. Chapter 2 establishes the

hypotheses and objectives of this investigation. The databases that were used to perform the experiments are detailed in chapter 3. Then, chapter 4 describes the methodology that has been used in each of the studies that form this work. The main results are described in chapter 5, being further discussed in chapter 6. Finally, the contributions of this dissertation, as well as the final conclusions, are detailed in chapter 7. The last sections are intended to complement this document by including: the papers of the compendium of publications (appendix A), the scientific achievements achieved during the Ph.D. (appendix B), and a brief summary in Spanish (appendix C).

Chapter 2

Hypothesis and objectives

As shown in the first chapter, EEG-based BCI technology still faces major issues that hinder its use for practical applications outside the laboratory. This doctoral dissertation was devoted to providing new methods and tools in order to solve some of these limitations, especially in the context of assistive systems for severely disabled people. The following section outlines the supporting hypotheses and main objectives of this work.

2.1 Hypothesis

This dissertation was focused on three of the limitations that affect ERP-based spellers: (1) synchronous control; (2) performance; and (3) software tools for BCI research. The hypotheses that guided each of the studies that compose the compendium are described in the following paragraphs.

Regarding the first limitation, we started from the following hypothesis: *the stimulation pattern used in ERP-based spellers elicits different types of brain activity that can be detected in the EEG*. Based on the work of Pinegger et al. (2015), we also assumed that *this brain activity can be used to provide a reliable detection of the user's control state over the BCI*. These hypotheses guided the first and third studies of the compendium.

Despite the fact that there were several works that studied the application of deep learning for EEG processing, and more specifically, ERP detection, the improvement achieved by these investigations was not as impressive as in other pattern recognition tasks (e.g., image recognition or natural language processing). This fact led to our next hypothesis: *more complex deep-learning architectures that*

take into account the specific spatiotemporal dynamics of the EEG could improve the performance of these models for BCI applications. This statement guided our research to increase the performance of asynchronous ERP-based spellers for assistive applications in the second and third studies of the compendium.

The third limitation that was addressed in this work is related to the available tools for BCI research. We identified a number of problems in current BCI platforms that were hindering the development and distribution of the proposed BCIs. This analysis led to the following hypothesis: *the development of a novel BCI platform could accelerate the research in this field.* This last hypothesis was the foundation of the fourth and last work of the compendium.

These statements are the main hypotheses that form the core of the present dissertation, which can be merged into the following global hypothesis:

> "*The development of novel signal processing methods and research tools can significantly advance the field of ERP-based BCIs, thereby bringing these systems closer to practical implementation for assistive applications.*"

2.2 Objectives

The main goal of this dissertation was to design, develop and test novel signal processing methodologies to improve the performance and usability of asynchronous ERP-based BCIs in an assistive context. In order to achieve this general objective, the following specific objectives were proposed:

I. To further characterize the EEG signal during control/non-control states in ERP-based spellers to provide new insights that could be later used to improve control state detection in these systems.

II. To optimize signal processing methods in ERP-based spellers to increase the performance of the control state detection stage using novel features based on the characterization of control/non-control states.

III. To increase the performance of ERP-based spellers for assistive applications using deep learning approaches.

IV. To develop a novel BCI platform to facilitate the design, implementation and distribution of custom BCI experiments and applications to be used throughout this investigation, allowing us to reach the previous objectives.

V. To disseminate the main results of this work in JCR indexed journals, and international/national conferences.

Chapter 3

Subjects and signals

In this chapter, the databases that were used in each of the studies that compose this dissertation are detailed. Section 3.1 introduces the row-column paradigm (RCP), which is the BCI control paradigm that was used in the first three studies of the compendium: Santamaría-Vázquez et al. (2019a, 2022, 2020b). Section 3.2 describes the clinical and demographic characteristics of the subjects that participated in the experiments. Finally, section 3.3 details the acquisition setups for each study, including EEG devices, sampling rate and electrode montages.

3.1 The row-column paradigm

As discussed in section 1.3.3, ERPs are the brain response to specific events. For instance, sudden visual stimuli provoke a series of characteristic EEG waveforms with different positive and negative peaks that can be revealed by averaging EEG epochs after the stimulus onset. Specifically, P300 evoked potentials are endogenous components elicited by the *oddball paradigm*, where the user has to detect an infrequent target stimuli among frequent non-target stimuli. This simple cognitive task triggers conscious responses that activate high-level areas of the cortex in frontal and parietal lobes whenever the user detects a target stimulus. This activity is reflected in the EEG with a characteristic waveform called P300 potential, which appears around 300 ms after the stimulus onset. Importantly, the P300 is not present for non-target stimulus (Sellers et al., 2012).

Different BCI paradigms take advantage of these differences in the EEG for target vs non-target stimuli. The RCP is probably the most known among them. In this paradigm, the system displays a matrix of commands, whose rows and

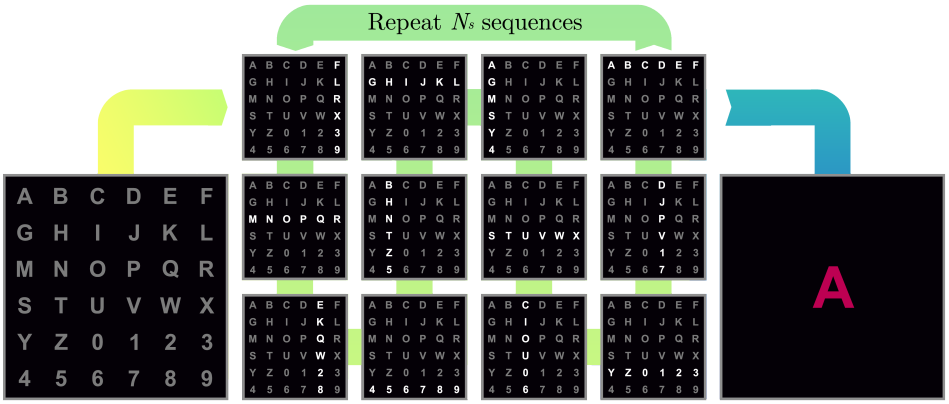

Figure 3.1: Representation a row-column paradigm (RCP) trial with a 6×6 matrix. When the trial starts, the system highlights all the rows and columns of the matrix. The process is repeated N_s times to increase the number of observations and improve the detection accuracy of the P300 potential. Finally, the BCI processes the EEG signal and decodes the command.

columns are highlighted randomly. To select a command, the user has to stare at the desired option, eliciting an ERP with the P300 component when a target stimulus is perceived. On the other hand, non-target stimuli does not trigger such response, only eliciting low-level responses in the early stages of the visual cortex related to unconscious processing. Finally, the system decodes the row and the column using signal processing algorithms to detect the P300 component and executes the corresponding command, providing feedback to the user. Figure 3.1 shows an schematic representation of the RCP operation. It should be noted that this paradigm is inherently synchronous: the system assumes by default that the user is always attending the stimuli and selects a command after every trial. In this regard, due to the low SNR of the ERPs, the control state detection task requires specific processing in order to detect the user's control state.

In order to configure the RCP operation, several parameters must be defined:

1) **Stimulus duration (SD).** This parameter defines the time that each row and column is highlighted. Its value is usually set between 75 to 200 ms. Values below 50 ms can prevent the detection of the stimulus by the user, leading to very low command decoding accuracies (Blankertz et al., 2006; Treder et al., 2011).

2) **Inter-stimulus interval (ISI).** This parameter defines the temporal separation between two consecutive stimuli. Typically, the ISI is set between 75

and 300 ms (Blankertz et al., 2006; Treder et al., 2011). A rather common practice is to let the value of the ISI change between stimulus according to a random uniform distribution to prevent users' anticipation (Martínez-Cagigal et al., 2017).

3) **Stimulus onset asynchrony (SOA).** The SOA is the sum of the SD and the ISI, representing the time between two consecutive onsets. It usually takes values in the range of 175–500 ms (Blankertz et al., 2006).

4) **Number of sequences (N_s).** In order to increase the accuracy of the system, the stimulation procedure is repeated N_s times to increase the accuracy of the system. Generally, calibration data is recorded with $10 \leq N_s \leq 15$ sequences, being reduced afterwards to increase the command selection speed (Martínez-Cagigal et al., 2017). The flashing order should be randomized for every sequence to assure one of the main principles to elicit suitable P300: an unexpected stimulation.

5) **Matrix dimensions.** The number of rows and columns of the matrix will determine the number of commands that can be allocated. One of the most useful features of the RCP is that the accuracy of the system is not compromised regardless the number of commands. However, it should be noted that a higher number of commands implies a higher selection time. Another important aspect is that there are few target stimuli that will elicit a P300 in comparison with non-target stimuli, leading to class-imbalanced datasets (e.g., 1:6 for a 6 x 6 matrix) that will affect the design of the signal processing algorithms.

Additional notation details that will be used along this document are: (1) a stimulus refer to the highlighting of a row or column; (2) a trial is a single selection with the system (e.g., one letter); and (3) a run is series of several trials that are selected to perform a specific task (e.g., write a word, browse a web).

Several studies proved that the RCP could be a reliable paradigm for assistive BCI applications (Martínez-Cagigal et al., 2017, 2019a). However, as stated in section 1.4, there is still room for improvement in terms of accuracy, speed and asynchronous management.

3.2 Subjects

Three independent databases were used during the course of this doctoral thesis:

1 **Asynchrony database.** This database was acquired to characterize the differences in the EEG during control *vs.* non-control states in our ERP-based speller with the RCP. Participants performed 120 trials, 60 control and 60 non-control, of 15 stimulation sequences with the proposed system. Signals were recorded using a 16-channel EEG cap. The database was recorded in two phases. The first 15 subjects (11 males, 4 females, mean age: 26.1 ± 2.3 years) were acquired for Santamaría-Vázquez et al. (2019a). Afterwards, this sample was extended for Santamaría-Vázquez et al. (2022) with 7 more subjects, making a total of 22 control subjects (CS) (15 males, 7 females, mean age: 24.7 ± 4.3 years). Table 3.1 shows the demographic information of subjects in this database.

2 **BCI Web Browser database.** This database was acquired to validate an RCP speller to help severely disabled people to surf the internet with an adapted BCI web browser (Martínez-Cagigal et al., 2017). The database consisted of 10 CS (6 males, 4 females, mean age: 24.8 ± 2.9) and 15 motor disabled subjects (MD) (10 males, 5 females, mean age: 42.7 ± 7.5 years) that performed different tasks with the web browser using the speller. Each subject performed 87.9 ± 7.3 control trials. It should be noted that some tasks contained free spelling moments, hence the variability in the number of trials per subject. Signals were recorded using an 8-channel EEG cap (Martínez-Cagigal et al., 2017). Demographic and clinical data of this database are shown in Table 3.2

3 **BCI Social Networks database.** This database contains data from a feasibility study to evaluate an assistive BCI application to use several social networks in a smartphone controlled through a RCP speller (Martínez-Cagigal et al., 2019a). The database contained data from 10 CS (8 males, 2 females, mean age: 26.10 ± 3.45) and 16 MD (10 males, 6 females, mean age: 45.5 ± 9.68 years). Each subject performed 63.4 ± 8.2 control trials using an 8-channel cap (Martínez-Cagigal et al., 2019a). Demographic and clinical data of this database are shown in Table 3.3.

Table 3.4 summarizes the characteristics of each database. All subjects gave their informed written consent for participating in the respective studies. The protocols were approved by the local ethics committee and complied with the declaration of Helsinki. MD subjects of "BCI Web browser study" and "BCI Social networks study" were recruited by the Spanish National Reference Centre on Disability and Dependence.

Table 3.1: Demographics of the Asynchrony database.

	User	Sex	Age
Control subjects	CS01	M	27
	CS02	M	26
	CS03	M	27
	CS04	M	30
	CS05	M	35
	CS06	M	26
	CS07	M	27
	CS08	M	27
	CS09	F	28
	CS10	F	26
	CS11	M	27
	CS12	F	25
	CS13	F	23
	CS14	M	26
	CS15	M	27
	CS16	F	19
	CS17	M	19
	CS18	M	19
	CS19	F	19
	CS20	M	19
	CS21	F	19
	CS22	M	26

MD: motor disabled subjects, M: male, F: female.

3.3 EEG recording setup

For the three databases, the data were recorded with a g.USBAmp (*g.tec Medical Engineering*, Austria), at a sampling rate of 256 Hz. The ground was placed in FPz and the EEG signal was referenced to the earlobe. The electrode positions were different across databases. For the Asynchrony database 16 EEG channels were recorded, whereas for the BCI Web Browser and BCI Social Networks we used 8 EEG channels. Figure 3.2 shows the EEG montages for the three databases according to the International System 10–10. For all the recordings, the participants were sat on a comfortable chair in front of a computer screen, keeping a minimum distance of 50 cm.

Table 3.2: Demographics and clinical data of the BCI Web Browser database.

	User	Sex	Age	Condition
Control subjects	CS23	M	26	Healthy
	CS24	M	24	Healthy
	CS25	F	25	Healthy
	CS26	M	26	Healthy
	CS27	F	23	Healthy
	CS28	F	28	Healthy
	CS29	M	30	Healthy
	CS30	M	20	Healthy
	CS31	M	24	Healthy
	CS32	F	22	Healthy
Motor disabled subjects	MD01	M	31	Multiple sclerosis
	MD02	M	43	Multiple sclerosis
	MD03	F	47	Multiple sclerosis
	MD04	M	56	Multiple sclerosis
	MD05	F	32	Multiple sclerosis
	MD06	M	35	Multiple sclerosis
	MD07	M	41	Multiple sclerosis
	MD08	F	49	Multiple sclerosis
	MD09	M	44	Multiple sclerosis
	MD10	F	41	Multiple sclerosis
	MD11	M	43	Multiple sclerosis
	MD12	M	44	Multiple sclerosis
	MD13	M	52	Multiple sclerosis
	MD14	F	38	Multiple sclerosis
	MD15	M	47	Multiple sclerosis

MD: motor disabled subjects, M: male, F: female.

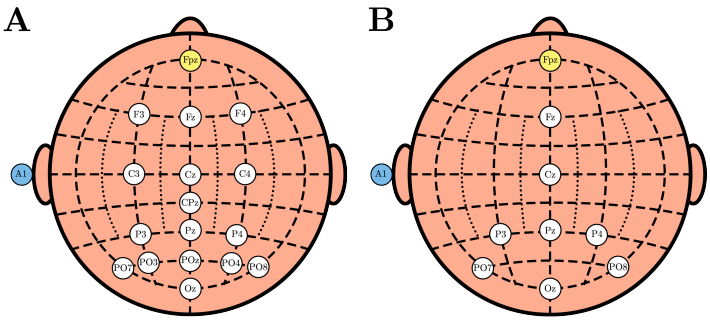

Figure 3.2: Electrode montages according to the International System 10–10: (A) Asynchrony database (Santamaría-Vázquez et al., 2019a, 2022); (B) BCI Web Browser and BCI Social Networks databases (Santamaría-Vázquez et al., 2020b). Ground and reference electrodes are marked in yellow and blue, respectively. Adapted from Martínez-Cagigal (2020).

Table 3.3: Demographics and clinical data of the BCI Social Networks database.

	Identifier	**Sex**	**Age**	**Condition**
Control subjects	CS33	M	25	Healthy
	CS34	M	25	Healthy
	CS35	M	24	Healthy
	CS36	M	25	Healthy
	CS37	M	25	Healthy
	CS38	M	32	Healthy
	CS39	M	24	Healthy
	CS40	M	25	Healthy
	CS41	F	23	Healthy
	CS42	F	33	Healthy
Motor disabled subjects	MD16	F	48	Stroke
	MD17	M	46	Spinal cord injury
	MD18	F	38	Friedreich's ataxia
	MD19	M	39	Spinal cord injury
	MD20	F	49	Friedreich's ataxia
	MD21	M	31	Cerebral palsy
	MD22	M	52	Cerebral palsy
	MD23	M	44	Friedreich's ataxia
	MD24	M	47	Cerebral palsy
	MD25	M	67	Cerebral palsy
	MD26	M	62	Muscular dystrophy
	MD27	M	47	Muscular dystrophy
	MD28	F	66	Friedreich's ataxia
	MD29	F	40	Friedreich's ataxia
	MD30	M	38	Spinal cord injury
	MD31	M	50	Spinal cord injury

MD: motor disabled subjects, M: male, F: female.

Table 3.4: Summary of each database.

Database	**CS**	**MD**	**Paradigm**	**SD**	**ISI**	**Channels**
Asynchrony	22	0	RCP	75	100	16
BCI Web Browser	10	15	RCP	62.5	$\mathcal{U}(125, 250)$	8
BCI Social Networks	10	16	RCP	62.5	$\mathcal{U}(125, 250)$	8

CS: control subjects; MD: motor disabled subjects; RCP: row-col paradigm; SD: stimulus duration in ms, ISI: inter-stimuli interval in ms; $\mathcal{U}$ randomized within uniform distribution.

Chapter 4

Methods

This chapter describes the methods that have been applied to conduct this doctoral dissertation. First, the signal processing methods are explained, including pre-processing in section 4.1, feature extraction in section 4.2, feature selection in section 4.3, and feature classification in section 4.4. It should be taken into account that in this document we only included the methods that were part of the proposed processing pipelines of each study, leaving out those that were used only for comparison purposes. After the signal processing methods, the performance metrics, statistical methodology and validation strategies that were used in this work are detailed in section 4.5. Finally, the methods and strategies to develop MEDUSA©, our novel BCI platform, are explained in section 4.6.

4.1 Pre-processing

Pre-processing is an important step in the analysis of EEG signals, as it helps to clean and prepare the raw data for further analysis. This is crucial because EEG signals often contain a significant amount of noise and artifacts, such as electrical interference, muscle activity, and eye movements. A proper pre-processing can help to remove or reduce these artifacts, improving the overall quality and reliability of EEG data. It can also help to extract relevant features from the EEG signal in further stages of the processing pipeline, such as frequency-based features or time-frequency representations, which can provide valuable information about the underlying brain activity. In this work, we applied two pre-processing techniques to enhance the EEG signals: frequency filtering and spatial filtering.

4.1.1 Frequency filtering

Although it is difficult to establish a unified criterion, it is widely accepted that the EEG carries information from the electrical activity of the brain in the range from 0.1 to 100 Hz approximately (see section 1.3.2). Therefore, frequencies outside this band should be removed before any analysis, making special emphasis in the elimination of the amplifier's low-frequency drifts in the 0 to 0.1 Hz band. The power line frequency (50 or 60 Hz) must also be removed to avoid this interference. Additionally, it should be taken into account that the power of the EEG features encoding the user's intentions in a BCI system is usually not spread across this entire spectrum. In that case, a narrower frequency filter can increase the SNR of the control signal by removing undesired EEG components. For instance, visual ERPs are low-frequency waveforms that have most of the power between 1 and 10 Hz, whereas SMRs are mainly detected from 10 to 30 Hz.

There are two types of digital filters that can be applied in this context: finite impulse response (FIR) and infinite impulse response (IIR) filters. In this work, we used FIR filters over IIR filters for: (1) linear phase response, making them particularly useful for applications where phase distortion is undesirable; (2) inherent mathematical stability, avoiding unwanted oscillations; and (3) non-feedback design, which simplifies the implementation and avoids compounding rounding errors their(Proakis, 2001). On the other hand, FIR filters require more computational resources to achieve the same level of performance as IIR filters. However, this increase in the computational cost is affordable in ERP-based BCIs, which have to process a trial every few seconds (Santamaría-Vázquez et al., 2019a, 2022, 2020b).

The output signal is calculated as the discrete convolution of the input signal with the filter coefficients, which determine the frequency response of the filter. The general equation for a FIR filter can be expressed as follows (Proakis, 2001):

$$y[n] = \sum_{k=0}^{N-1} h[k] \cdot x[n-k], \tag{4.1}$$

where $y[n]$ is the output signal at time n, $x[n]$ is the input signal at time n, $h[k]$ are the filter coefficients, and N is the number of coefficients in the filter.

In this dissertation, EEG signals were pre-processed by applying FIR band-pass filters with Hamming window and order 1000. The frequency ranges depended on the specific features that were extracted for each analysis. These bands are detailed in following sections, where the different processing pipelines are described.

4.1.2 Spatial filtering

Spatial filters can be used to minimize the volume conduction problem, remove noisy artefacts, or increase the SNR of localized control signals, among others. In this work, we applied a common average reference (CAR) filter to enhance the detection of ERPs. The CAR filter provides an estimate for the ideal reference-independent potentials at each recording location. Importantly, this estimation is ideal for montages covering a surface that contains all current within the head volume. Although the entire head surface cannot be covered (e.g., face and neck), the error of the approximation is reduced as the number of electrodes grow (Srinivasan, 2012). This method also allows to reduce artifacts common to all electrodes, such as power interference. Since the recorded EEG signal was already referenced to the ear lobe in our experiments (see section 3.3), the CAR filter is calculated by subtracting to each channel the mean potential of all the electrodes. The general equation for the CAR filter can be expressed as follows (Srinivasan, 2012):

$$\boldsymbol{x}_c = \boldsymbol{x}_c - \frac{1}{N_c}\sum_{i=1}^{N_c}\boldsymbol{x}_i, \tag{4.2}$$

where $\boldsymbol{x}_c$ is the signal for the channel c, and N_c is the number of channels.

4.2 Feature extraction

In machine learning, features are measurable properties or characteristics of a phenomenon being observed that serve as the input data to the classification model. Features can be any type of data, such as numeric, categorical, or binary values. In the EEG domain, the most basic features are the amplitude values of the signal at each location and temporal point. Then, these values can be transformed, using a wide range of mathematical methods, to obtain more meaningful information for the analysis. Overall, the choice of features is crucial in the development of EEG models, as it can have a significant impact on the performance and accuracy of the method. Proper feature engineering can help to improve the predictive power of an EEG model and make it more effective at solving specific problems, such as ERP detection or MI classification.

Taking into account the strict temporal requirements of the EEG processing stage in a BCI system, the aim is to maintain this stage as simple as possible using computationally efficient algorithms. In this investigation we used temporal, spectral and correlation features for the control state detection and command

decoding tasks of our ERP-based BCI.

4.2.1 Time-based features

Time-based features are directly related with the amplitude values of the EEG over time. Thus, they represent the most primary source of information of this signal. In Santamaría-Vázquez et al. (2019a, 2022, 2020b), we used these features to implement different algorithms for the control state detection and command decoding stages of our ERP-based BCI, analyzing the EEG signal corresponding to the stimulation events of the RCP. The process that was followed in this studies is detailed in the following paragraphs.

First, we decimated the signal to reduce its dimensionality, decreasing the sampling rate from f_s to f_d. Noteworthy, a low-pass filter must be applied before to avoid aliasing effects with a cutt-off frequency of $f_d/2$ in accordance to the Nyquist–Shannon theorem (Proakis, 2001).

After decimation, the EEG values that correspond to each flash of the RCP are extracted using a temporal window that typically lasts from the stimulus onset (i.e. $t = 0$ ms) to a point in time that captures the entire P300 component (e.g. 600-1000 ms). These chunks of signal are known as epochs. Then, EEG epochs are normalized using a z-score baseline normalization, based on a previous reference time window. This mathematical operation can be expressed as follows:

$$\boldsymbol{x} = \frac{\boldsymbol{x} - \mu_{\boldsymbol{x}_R}}{\sigma_{\boldsymbol{x}_R}}, \tag{4.3}$$

where $\boldsymbol{x}_R$ is the EEG signal of the reference window, and μ and σ are the mean and standard deviation, respectively.

Finally, our feature arrays are EEG epochs with shape $\boldsymbol{x} \in \mathbb{R}^{N_s, N_c}$, where N_s is the number of samples in the window and N_c is the number of channels. This array was the input of the deep-learning models (i.e., EEG-Inception, EEGNet and DeepConvNet, CNN-BLSTM) (Santamaría-Vázquez et al., 2019a, 2022, 2020b). For the machine-learning models (i.e., LDA), the channels were concatenated to obtain a feature vector $\boldsymbol{x} \in \mathbb{R}^{1, N_s \times N_c}$ as required by these methods (Santamaría-Vázquez et al., 2019a, 2020b).

4.2.2 Frequency-based features

As explained in 1.3.2, spectral analysis is a common technique used to analyze the frequency content of the EEG. The signal is decomposed into its frequency

components using techniques based on Fourier transform or wavelet transform (Tantum, 2021). The resulting frequency spectrum gives the power of the EEG at each frequency, which can be used to extract features that capture the underlying neural activity.

There are many different features that can be extracted from the frequency spectrum. The most typical are band powers, which quantify the power of EEG signals within specific frequency bands. For example, alpha power is often used as a measure of relaxation, while beta power is often associated with cognitive load. The use of band-power ratios is also extended. For instance, the ratio of alpha power to beta power (theta/beta ratio) is often correlated with performance in various cognitive abilities, such as attentional control (Angelidis et al., 2018). Other metrics, such as spectral entropy, have been proposed to study different conditions (Abásolo et al., 2006).

In Santamaría-Vázquez et al. (2019a), we used a frequency-based feature designed to detect the user's control state in ERP-based spellers. The feature was a band-power ratio directly derived from the PSD of the EEG to detect the SSVEPs that are elicited by non-target stimuli in the RCP (Santamaría-Vázquez et al., 2019a).

Firstly, all channels of the signal corresponding to an entire trial (not to a single stimulus) were concatenated in a single vector to increase the spectral resolution. Afterwards, the PSD of this vector was estimated using the Welch's method (Welch, 1967). From the PSD, the feature is calculated as the difference between the mean value in a narrow range and the mean value in a wide range, as shown in see Figure 4.1(a). Both ranges are centered in the stimulation frequency of the ERP-based speller. Formally:

$$x_2 = \frac{1}{bw_2}\int_{f_{st}-bw_2/2}^{f_{st}+bw_2/2} S(f)df - \frac{1}{bw_1}\int_{f_{st}-bw_1/2}^{f_{st}+bw_1/2} S(f)df, \tag{4.4}$$

where $S(f)$ is the PSD, and bw_1 and bw_2 are fixed to 2 Hz and 0.1 Hz, respectively (Santamaría-Vázquez et al., 2019a). These values were chosen to have enough points of the PSD within these bands, making the method more robust against noise.

4.2.3 Correlation-based features

Correlation-based features measure the degree of similarity between two datasets. In BCI, these features are widely used to detect SSVEPs. As explained in 1.3.3,

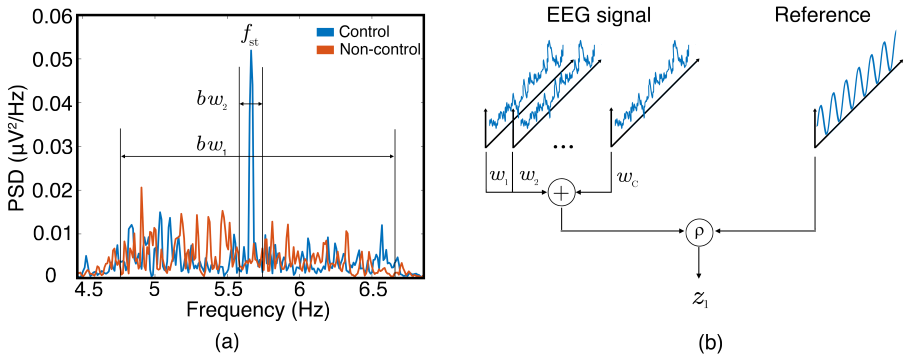

Figure 4.1: (a) Representation of the frequency-based feature, including the parameters f_{st}, bw_1, bw_2. The graph shows the averaged power spectral density (PSD) of control and non-control trials of the Asynchrony database. (b) Schematic representation of the canonical correlation analysis (CCA) used to extract the correlation-based. The EEG signal represents one trial of dimensions $C \times N$, where C is the number of channels and N the trial length in samples. The reference signal is an ideal sinusoid $sin(2\pi \cdot f_{st} \cdot t)$ of dimensions $1 \times N$.

SSVEPs are the result of repetitive stimulation, in which individual VEPs are overlapped producing a constant oscillation at the stimulation frequency that resembles a sinusoidal wave (Luck, 2014). Therefore, the correlation between the EEG during the stimulation and a sine at the stimulation frequency is high, allowing to differentiate between different commands in SSVEP-based BCIs (Bin et al., 2009; Lin et al., 2006).

In Santamaría-Vázquez et al. (2019a), we used a correlation-based feature to identify the user's control state in ERP-based spellers. The goal was to detect the SSVEP elicited by non-target stimuli of the system. Concretely, we used canonical correlation analysis (CCA), a multivariate statistical technique that aims to identify the relationships between two multidimensional datasets, denoted as $\boldsymbol{X}$ and $\boldsymbol{Y}$ (Krzanowski, 2000). It finds the optimal linear combinations $\boldsymbol{x} = \boldsymbol{w}_x^T * \boldsymbol{X}$, $\boldsymbol{y} = \boldsymbol{w}_y^T * \boldsymbol{Y}$ of the datasets, such that the correlation between $\boldsymbol{x}$ and $\boldsymbol{y}$ is maximized. In our study, $\boldsymbol{X}$ represents the EEG signal of a single trial with an ERP-based speller, with dimensions $N \times C$, where N is the length of the trial and C is the number of channels. $\boldsymbol{Y}$ represents the reference signal, which is a sine wave at the stimulation frequency of the RCP f_{st}, and has dimensions $N \times 1$. The calculated feature corresponds to the correlation coefficient between the trial signal and the reference calculated using the following equation:

$$\rho_c = \max_{\boldsymbol{w_x}, \boldsymbol{w_y}} \frac{\boldsymbol{w_x}^T \Sigma_{XY} \boldsymbol{w_y}}{\sqrt{\boldsymbol{w_x}^T \Sigma_{XX} \boldsymbol{w_x}} \sqrt{\boldsymbol{w_y}^T \Sigma_{YY} \boldsymbol{w_y}}}, \tag{4.5}$$

where ρ_c is the canonical correlation coefficient, $\boldsymbol{w_x}$ and $\boldsymbol{w_y}$ are the canonical weights for the two sets of variables, and Σ_{XY}, Σ_{XX}, and Σ_{XY} are the covariance matrices of the two sets of variables. Figure 4.1(b) shows a schematic representation of the CCA method.

CCA had previously been applied in SSVEP-based BCIs, but, to the best of our knowledge, Santamaría-Vázquez et al. (2019a) was the first study that used it for control state detection in ERP-based spellers.

4.3 Feature selection

The feature selection stage is designed to optimize the data passed to classification models by eliminating redundant and irrelevant information from the dataset, thus retaining only the most important features. This not only reduces the number of features, but also improves the model's resistance to overfitting and increases its efficiency (Jobson, 1991). In this dissertation, we only used one feature selection method as part of the command decoding pipeline in Santamaría-Vázquez et al. (2019a): stepwise regression.

4.3.1 Stepwise regression

Stepwise (SW) regression is a widely used feature selection algorithm in ERP-based BCIs, as described in various studies such as Krusienski et al. (2008) or (Sellers et al., 2012). The SW algorithm follows a sequential process, where it either adds or removes a single feature at each step. The process to add new features is called forward selection (FS), whereas the process eliminate a features is named backward elimination (BE) (Jobson, 1991).

The SW algorithm starts from an empty dataset. First, it applies FS by testing the significance of adding each feature separately, according to a partial F-statistic. The most statistically significant feature, provided its p–value is less than a preset threshold p_{in}, is then added to the model. After each new feature is added, the algorithm performs BS, testing the significance of each included feature again. The less significant feature, given that the p–value is greater than another preset threshold p_{out}, is then removed from the model. This sequential combination of FS and BE procedures is repeated until one of three conditions is met: (1) there are

no features that meet the p_{in} and p_{out} criteria; (2) there are no remaining features; or (3) the number of included features reach a predefined limit. It is important to note that in order to prevent an infinite loop, the p_{in} threshold should be at least as small as p_{out} (Jobson, 1991).

In Santamaría-Vázquez et al. (2019a), SW regression was used with the following criteria: $p_{in} = 0.10$, $p_{out} = 0.15$ and a maximum of 60 features. These values were selected in accordance to previous studies (Krusienski et al., 2008).

4.4 Feature classification

In this dissertation, we used machine learning algorithms in the feature classification stage to decide the presence/absence of the pattern we are looking for in the EEG and decode the user's intentions. Thus, it is the final step of the signal processing pipeline in a BCI system.

Machine learning is a subfield of artificial intelligence that involves the development of algorithms and models that allow computer systems to learn and improve their performance on a specific task without explicit programming. These algorithms and models are designed to recognize patterns and relationships in data, making predictions or decisions based on that analysis (Bishop and Nasrabadi, 2006).

There are several types of machine learning, including supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning (Bishop and Nasrabadi, 2006). In this dissertation, we only applied supervised learning, which is the most robust approach for EEG classification tasks.

In supervised learning, the model infers a discriminant function to make predictions based on a labeled dataset, which includes a collection of observations from the phenomena that need to be detected, along with the corresponding correct output for each observation (Bishop and Nasrabadi, 2006). The model is then trained by minimizing the error between its predictions and the correct output.

In this work, the models were used for two tasks: (1) control state detection to provide an asynchronous control of our ERP-based speller; and (2) command decoding to detect the user's intentions. The training data consisted of labeled EEG recordings while the user was using the system. For the control state detection, we recorded the brain activity of the subjects while they were using the system and while they were performing other tasks, labeling the control trials as positive, otherwise as negative. For the command decoding task, subjects were asked to spell predefined words using the system, labeling the stimuli that corresponds

to the target character as positive, otherwise as negative. Once the models were trained, the system could classify new online or offline data (Santamaría-Vázquez et al., 2019a, 2022, 2020b).

In the following subsections, the classifiers that were used to implement the proposed methods in this dissertation are described: linear discriminant analysis and EEG-Inception.

4.4.1 Linear Discriminant Analysis

Linear discriminant analysis (LDA) is a method to separate observations from different classes using a linear combination of features. In this dissertation, we used this method as a classification technique, although it can be applied for dimensionality reduction too. In the training stage, LDA calculates the linear combination of features that best separates the classes by maximizing the separation between-class mean while simultaneously minimizing the within-class variance. In the testing stage, this linear combination is used to make predictions with new data.

For a binary classification problem like the one being considered, the weight vector can be found by solving the following optimization problem (Bishop and Nasrabadi, 2006):

$$\max J(\boldsymbol{w}) = \frac{\boldsymbol{w}^T \boldsymbol{S}_B \boldsymbol{w}}{\boldsymbol{w}^T \boldsymbol{S}_W \boldsymbol{w}}, \tag{4.6}$$

$$\boldsymbol{S}_B = (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T, \tag{4.7}$$

$$\boldsymbol{S}_W = \sum_{n \in \mathcal{C}_1} (\boldsymbol{X}_n - \boldsymbol{\mu}_1)(\boldsymbol{X}_n - \boldsymbol{\mu}_1)^T + \sum_{n \in \mathcal{C}_2} (\boldsymbol{X}_n - \boldsymbol{\mu}_2)(\boldsymbol{X}_n - \boldsymbol{\mu}_2)^T, \tag{4.8}$$

where $\boldsymbol{S}_B$ is the between-class matrix, $\boldsymbol{S}_W$ the within-class matrix, and $\boldsymbol{\mu_i}$ the mean of class $i \in \mathcal{C}_{\{1,2\}}$. Given these equations, the optimal weights are:

$$\boldsymbol{w} \propto \boldsymbol{S}_W^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1). \tag{4.9}$$

Once the optimal vector w is found, it can be used to define a linear decision boundary, which can be written as $w^T x + b = 0$, where b is a bias term. Generally, this term is chosen as the hyperplane between the class means: $b = \boldsymbol{w}^T \cdot \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)$. New data points can then be classified by evaluating their position relative to this decision boundary. Points on one side of the boundary belong to one class, while points on the other side belong to the other class.

In order to reduce the over-fitting effect and increase the robustness of the classifier to outliers, there are different regularization techniques that can be applied. A widely used method is to introduce shrinkage to improve the estimation of covariance matrices. This technique increases the generalization performance of the classifier when the number of training samples is small. Generally, the regularized covariance estimator is defined as (Ledoit and Wolf, 2004):

$$\boldsymbol{S^*} = (1 - \lambda)S + \lambda\mu I, \tag{4.10}$$

$$\mu = \frac{Tr(S)}{N_f}, \tag{4.11}$$

where $\lambda \in (0, 1)$ is the shrinkage parameter, S is the sample covariance, I is the identity matrix, Tr represents the sum of the diagonal elements, and N_f is the number of features. In this dissertation, the shrinkage parameter was calculated using the Ledoit and Wolf formula as implemented in the scikit-learn package (Pedregosa et al., 2011). When this technique was applied, we denote to the classifier as regularized LDA (rLDA) (Santamaría-Vázquez et al., 2020b).

Finally, it is worth noting that, for the command decoding task, an additional step is needed to decode the command. For each trial, we have $(R+C) \times N_s$ EEG epochs (i.e., observations), being R the number of rows of the RCP matrix, C the number of columns, and N_s the number of sequences. Each epoch corresponds to one stimulus (i.e., highlight of a row or a column). The LDA algorithm then decides, for each observation, whether it contains an ERP with P300 potential (i.e., target epoch) or not. Thus, in order to decide the command, the scores for each row and column are averaged, selecting the pair of them with highest probability for the target class.

4.4.2 EEG-Inception

EEG-Inception is a novel CNN that was proposed in Santamaría-Vázquez et al. (2020b) for EEG classification tasks. In comparison to a machine learning classifier, such as LDA, deep learning models incorporate an additional feature extraction stage embedded in the architecture that is automatically optimized during training (Lecun et al., 2015). Concretely, EEG-Inception is a model composed of multiple layers that extract hierarchical representations of the input data. In general, the first layers of the network learn to extract simple, low-level features from the input data. However, as the data passes through the network and it is processed by deeper layers, the features become increasingly abstract and complex,

capturing high-level information about the signal.

EEG-Inception was inspired by previous works on the computer vision field, incorporating several concepts from this domain, such as Inception modules and depthwise convolutions, to improve the processing of EEG signals (Santamaría-Vázquez et al., 2020b).

Inception modules were proposed by Szegedy et al. (2015) for image classification. This structure uses multiple convolutional layers in parallel with different kernel sizes to perform multiscale analysis of the input signal. This approach has been shown to be effective at extracting rich feature maps with relatively low computational cost, improving performance while maintaining a reasonable training and evaluation time. Inception modules are particularly well-suited for analyzing EEG signals, which are composed of transient and oscillatory patterns of various temporal lengths that reflect brain activity (Santamaría-Vázquez et al., 2020b).

Depthwise convolutions were introduce to allow the model to learn more targeted and specialized features in a CNN (Chollet, 2015). Standard convolutional layers filter the input data using the same set of weights for all the feature maps. Depthwise convolutions, on the other hand, apply a different set of weights to each feature map. Thus, the model processes each map separately, rather than combining them as in standard convolutions. In EEG-Inception, depthwise convolutions enable the learning of specific features for each feature map extracted by the Inception modules. This is a key feature for the model, as it will be revealed bellow (Santamaría-Vázquez et al., 2020b).

Additionally, EEG-Inception incorporates standard methods to increase the performance of the model for EEG processing, including batch normalization to normalize the feature maps (Ioffe and Szegedy, 2015), non-linear activation functions (Goodfellow et al., 2016), and dropout regularization (Srivastava et al., 2014) to prevent overfitting.

The architecture of EEG-Inception is shown in Figure 4.2, whereas Table 4.1 provides further details about the hyperparameters of each layer. The input expected by the model is the time-based feature array as described in section 4.2.1 with decimation to 128 Hz sampling rate and eight channels Santamaría-Vázquez et al. (2020b). Then, the signal is processed by the model, which is organized in three main blocks, each one with a different purpose. The first module performs a first temporal and spatial analysis of the EEG to extract low-level features in both domains. The second module performs a second analysis that combines the temporal and spatial information to extract high-level features. Finally, the output block is responsible for the final classification of the input signal. In the following

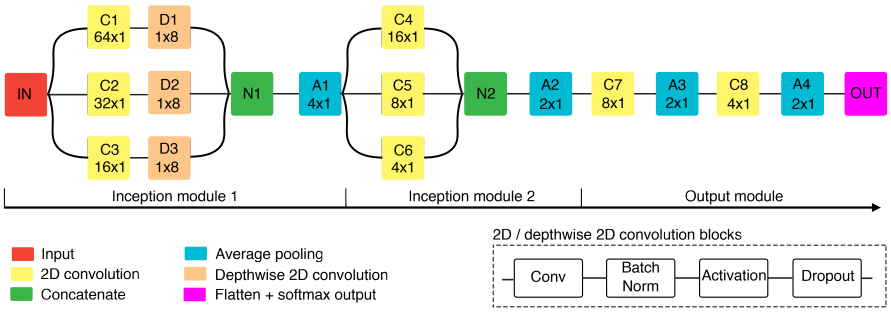

Figure 4.2: Schematic representation of EEG-Inception. 2D convolution blocks and depthwise 2D convolution blocks include batch normalization, non-linear activation and dropout regularization. The kernel size is displayed for convolutional and average pooling layers.

sections, the three modules of EEG-Inception are explained in detail:

Inception module 1

The first Inception module is designed to perform a first stage of temporal and spatial analysis. The initial 2D convolutions analyze the input signal in the time domain using three branches, each one with a different temporal scale. The temporal scales are defined by the kernel sizes of the 2D convolutional layers: C1, C2 and C3. The kernels sizes of these layers are 64×1, 32×1 and 16×1, respectively. Therefore, given that the sampling rate of the input is 128 Hz, these kernels correspond to temporal windows of 500 ms, 250 ms and 125 ms. Following these layers, D1, D2 and D3 process the signal in the spatial domain using depthwise convolutions. In this architecture, this type of convolutions allow the model to learn optimal spatial filters (i.e., EEG channels weights) for each temporal feature map extracted by previous layers (Santamaría-Vázquez et al., 2020b). Finally, the concatenation layer N1 merges the output features from D1, D2 and D3. Finally, average pooling is applied for dimensionality reduction.

Inception module 2

The second module has a similar structure to the previous one. It is formed by three branches that process the feature maps in the same temporal scales: 500 ms, 250 ms and 125 ms. It should be noted that, after the previous average pooling layer (A1), the kernel sizes that correspond to these temporal scales are 16×1, 8×1 and 4×1. As in the first module, the outputs of convolutional layers are

concatenated and average pooling is applied. The purpose of this module is to extract higher level features, already taking into account the temporal and spatial information, which was mixed in the previous module.

Output module

The final convolutional layers of the model are responsible for identifying the most relevant patterns for the classification task and condensing this information into a small number of features. The dimensionality is reduced by progressively decreasing the number of filters and applying average pooling layers to avoid overfitting. In fact, only 24 features are passed to the final classification layer. Finally, the softmax output performs the binary classification, estimating the probability of each class (Santamaría-Vázquez et al., 2020b).

Hyperparameter choice and training process

The selection of hyperparameters in deep learning models is critical for achieving good results. For EEG-Inception, we automatically optimized the learning rate, activation function, and dropout rate due to their significant impact on model performance (Santamaría-Vázquez et al., 2020b). Unfortunately, an automatic optimization of all hyperparameters is impractical due to the cost in time and computational resources of this process. Thus, the remaining hyperparameters (e.g., number of layers, number of branches in Inception modules, number of filters, kernel sizes, and pooling sizes) were set based on our experience. The results of the automatic optimization process are detailed in section 5.3.1.

Regarding the training process, the model wights were optimized using the Adam algorithm with default hyperparameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$ (Kingma and Ba, 2015), the categorical cross entropy loss function (Zhang and Sabuncu, 2018), and a mini-batch size of 1024. The model was trained for a maximum of 500 epochs with early stopping if the loss on the validation set did not improve for 10 consecutive epochs. This early stopping mechanism speeds up training and prevents overfitting (Santamaría-Vázquez et al., 2020b).

Table 4.1: EEG-Inception architecture details

Block	Type	Filters	Depth	Kernel	Padding	Output shape	Connected to	Role
IN	Input	–	–	–	–	$128 \times 8 \times 1$	C1, C2, C3	Input
C1	Conv2D	8	–	64×1	Same	$128 \times 8 \times 8$	D1	Temporal analysis
D1	DepthwiseConv2D	–	2	1×8	Valid	$128 \times 1 \times 16$	N1	Spatial analysis
C2	Conv2D	8	–	32×1	Same	$128 \times 8 \times 8$	D2	Temporal analysis
D2	DepthwiseConv2D	–	2	1×8	Valid	$128 \times 1 \times 16$	N1	Spatial analysis
C3	Conv2D	8	–	16×1	Same	$128 \times 8 \times 8$	D3	Temporal analysis
D3	DepthwiseConv2D	–	2	1×8	Valid	$128 \times 1 \times 16$	N1	Spatial analysis
N1	Concatenate	–	–	–	–	$128 \times 1 \times 48$	A1	Concatenation
A1	AveragePooling2D	–	–	4×1	–	$32 \times 1 \times 48$	C4, C5, C6	Concatenation
C4	Conv2D	8	–	16×1	Same	$32 \times 1 \times 8$	N2	Temporal analysis
C5	Conv2D	8	–	8×1	Same	$32 \times 1 \times 8$	N2	Temporal analysis
C6	Conv2D	8	–	4×1	Same	$32 \times 1 \times 8$	N2	Temporal analysis
N1	Concatenate	–	–	–	–	$32 \times 1 \times 24$	A2	Concatenation
A2	AveragePooling2D	–	–	2×1	–	$16 \times 1 \times 24$	C7	Dimension reduction
C7	Conv2D	12	–	8×1	Same	$16 \times 1 \times 12$	A3	Temporal analysis
A3	AveragePooling2D	–	–	2×1	–	$8 \times 1 \times 12$	C8	Dimension reduction
C8	Conv2D	6	–	4×1	Same	$8 \times 1 \times 6$	A4	Temporal analysis
A4	AveragePooling2D	–	–	2×1	–	$4 \times 1 \times 6$	C7	Dimension reduction
OUT	Dense	–	–	–	–	2	–	Softmax output

Column "Type" describes the class used to implement each block in Keras framework. It should be taken into account that this implementation may vary across different frameworks. All convolutional blocks (i.e., Conv2D and DethpwiseConv2D) include batch normalization, activation and dropout regularization. The model has 15154 parameters, of which 14926 are fitted during training.

4.5 Validation

4.5.1 Performance metrics

When an ERP-based speller is validated, there are several ways of measuring its performance. For this reason, different metrics were needed to provide a comprehensive evaluation of the system:

1) **Accuracy.** This metric represents the percentage of correctly classified trials:

$$\text{Acc} = \frac{N_{ct}}{N_t} \times 100, \tag{4.12}$$

where, N_{ct} is the number of correctly classified trials, and N_t the total number of trials. In this case, we used this metric to evaluate different aspects of the proposed system. For this reason, several suffixes are used throughout this document to identify these aspects and avoid confusion. Acc_{csd} denotes the accuracy of the control state detection stage, i.e., the percentage of trials in which the user's control state was correctly identified. As there are two possible states (control or non-control), the chance level is 50%. Acc_{cmd} represents the accuracy of the command decoding stage, i.e., percentage of correct selections (letters or commands). In this case, the chance level is $\frac{1}{N_c} \times 100$, where N_c is the number of possible commands in the RCP matrix. Finally, Acc_{ovr} is the accuracy of the overall asynchronous system, including the control state detection and command decoding stages. To calculate Acc_{ovr}, a trial is considered correct only if both the control state and the command are accurately detected.

2) **Positive predictive value (PPV).** This metric represents the probability of having an actual positive case, given a positive prediction:

$$\text{PPV} = \frac{N_{tp}}{N_{tp} + N_{fp}} \times 100, \tag{4.13}$$

where, N_{tp} is the number of true positives, and N_{fp} the number of false positives.

3) **Negative predictive value (NPV).** This metric represents the probability of having an actual negative case, given a negative prediction:

$$\text{NPV} = \frac{N_{tn}}{N_{tn} + N_{fn}} \times 100, \tag{4.14}$$

where, N_{tn} is the number of true negatives, and N_{fn} the number of false negatives.

4) **True positive rate (TPR).** This metric, also known as sensitivity, represents the probability of a positive result, conditioned on the trial truly being positive:

$$\text{TPR} = \frac{N_{tp}}{N_{tp} + N_{fn}} \times 100, \tag{4.15}$$

where, N_{tp} is the number of true positives, and N_{fn} the number of false negatives.

5) **True negative rate (TNR).** This metric, also known as specificity, represents the probability of a negative result, conditioned on the trial truly being negative:

$$\text{TNR} = \frac{N_{tn}}{N_{tp} + N_{fp}} \times 100, \tag{4.16}$$

where, N_{tn} is the number of true negatives, and N_{fp} the number of false positives.

6) **Information transfer rate (ITR).** The ITR takes into account not only the command decoding accuracy, but also the speed of the system. Using the Shannon's information theory, the ITR estimates the number of bits per minute that the system is able to convey (Wolpaw et al., 2002):

$$\text{ITR} = \frac{1}{T}\left(\log_2 N + P\log_2 P + (1-P)\log_2 \frac{1-P}{N-1}\right), \tag{4.17}$$

where N is the number of trial selections, P is the selection accuracy (i.e., Acc_{cmd} for synchronous systems, and Acc_{ovr} for asynchronous systems), and T_s is the average duration of a trial in seconds.

4.5.2 Statistical analysis

Statistical tests are applied to make a fair comparison between two groups of results and extract the correct conclusions from the analysis. These tests evaluate the evidence provided by the data against a null hypothesis H_0. The p-value returned by the test represents the probability of getting the observation assuming that H_0 is true. If this probability is lower than a certain threshold α, it is considered that H_0 is false. A typical value for α is 0.05, representing a probability of 5%. When these tests are applied for group comparison, H_0 assumes that there are no

differences between the two groups of data. If H_0 is rejected, then we can talk about significant differences (Narsky and Porter, 2013).

In this dissertation, we applied two continuous, univariate, non-parametric statistical tests. In this regard, we used non-parametric tests because variables under study did not passed normality and/or homoscedasticity tests (Jobson, 1991). First, the Wilcoxon signed-rank test was used to assess paired comparisons (i.e., dependent samples that came from the same subjects) to analyze the statistical differences between the proposed pipelines and the previous methods in Santamaría-Vázquez et al. (2019a, 2022, 2020b). On the other hand, the Mann-Whitney U test, also known as the Wilcoxon rank-sum test, was used for unpaired comparisons (i.e., independent samples that came from different subjects) (Santamaría-Vázquez et al., 2022). When multiple tests were performed (e.g., comparisons among several models, for different number of sequences, etc), multiple testing correction was applied to correct the false discovery rate (FDR) using the Benjamini-Hochberg approach (Benjamini and Hochberg, 1995).

4.5.3 Cross-validation

Fundamentally, cross-validation techniques are used to assess the generalization ability of the results achieved by a classification algorithm. In other words, cross-validation allows to evaluate if a model can generalize to an independent dataset, which is not assured if the training samples are not representative (Bishop and Nasrabadi, 2006).

In this dissertation, we used leave-one-out (LOO) cross validation in Santamaría-Vázquez et al. (2019a, 2022, 2020b) to test the models in the different classification tasks. This method is a particularization of the k-fold cross-validation, keeping in each iteration one sample (or subject) for testing, whereas the rest of the dataset is used for training. This process is repeated for every observation in the dataset, so that each sample is used once as the test set. Then, the classification results are used to calculate the performance of the model using different metrics (e.g., accuracy). This technique provides an adequate estimate of the generalization ability of the model when the model is applied to new, unseen data without the need of a separate validation set (Bishop and Nasrabadi, 2006).

LOO can be computationally expensive, especially for large datasets, but it is a robust technique for evaluating models with small sample sizes. It is also useful when the dataset is unbalanced, having different number of observations for each class (Bishop and Nasrabadi, 2006).

4.6 BCI platform development

As mentioned in section 1.4.3, the lack of suitable software tools to achieve our primary research objectives led us to the development of a novel BCI platform: MEDUSA©. The goal was not only provide ourselves with the tools to conduct more efficiently this and other side projects that were developed during the dissertation period, but to make these tools available for the whole BCI community in the hope of making an difference in the progress of this field. In this section, we explain the strategies and software development methodologies that we used to achieve this goal.

4.6.1 Components

MEDUSA© is a software ecosystem for the development of BCI and neuroscience experiments, consisting of two independent components: MEDUSA© Kernel and MEDUSA© Platform.

MEDUSA© Kernel is a Python package designed for the analysis of brain signals. It offers a variety of methods, including advanced signal processing techniques, deep learning architectures, BCI models and other high and low-level analyses. The package also provides tools for handling bio-signals (e.g., EEG, MEG or EMG), saving experimental data and creating processing pipelines.

MEDUSA© Platform is a desktop application, also written in Python, that offers high-level functions for conducting BCI and neuroscience experiments. It includes a user-friendly graphical user interface (GUI) supported by advanced signal acquisition functions and real-time charts. One of its key features is the ability to create "apps", which are independent implementations of experiments or paradigms. All of these functions rely on MEDUSA© Kernel for the signal processing.

The features of both components are detailed in section 5.5, where the results of this work are explained.

4.6.2 Design principles

The design of MEDUSA© was guided by three principles:

- **Modularity.** MEDUSA© is composed of autonomous structures organized at various levels of abstraction that are connected through simple communication protocols. This design allows for quick fixes or upgrades to be made

without affecting the rest of the system. For example, the MEDUSA© Kernel has different interfaces for low-level and high-level functions with specific implementations that promote the independence between them. Similarly, the MEDUSA© Platform is designed to allow the creation of new experimental protocols on demand using specific components, known as apps, which are independent of the real-time acquisition and visualization stages.

- **Flexibility.** MEDUSA© was specifically designed to be a research tool, with an architecture that enables rapid experimentation with new signal processing methods and feedback paradigms. We placed a strong emphasis on providing thorough documentation and code comments, as well as examples and tutorials, to help users understand the platform's operation and how to develop new apps.

- **Scalability.** The design of MEDUSA© allows it to update its capabilities over time without requiring changes to unrelated parts of code, thanks to the use of standardized meta-classes. This is especially useful in a research setting, as it enables the software to adapt to the latest developments in the BCI field, such as new signal processing algorithms or BCI paradigms.

4.6.3 Implemented in Python

MEDUSA© was created using Python, a popular and open-source programming language that is commonly used in both research and industry due to its ease of use and open-source nature. In comparison to other languages like C, C++, or Java, Python makes it simpler to develop complex programs thanks to its high-level syntax, at the expense of a decrease in performance. This trade-off is especially beneficial in BCI research settings, where flexibility is a key aspect as new methods and experiments are frequently being developed. Python also has a large community that develops a variety of specialized tools and libraries. MEDUSA© takes advantage of packages like SciPy, Numpy, Scikit-learn, and Tensorflow, which were developed to provide state-of-the-art data processing, machine learning, and deep learning capabilities (Harris et al., 2020; Virtanen et al., 2020). This gives MEDUSA© an advantage over other neurotechnology platforms developed in C++, such as BCI2000 and OpenVibe, allowing us to easily incorporate and use the latest developments in these areas into the software workflow to improve different BCI paradigms.

4.6.4 Open-source community-oriented philosophy

MEDUSA© is an open-source software suite that is available under a Creative Commons Attribution-NonCommercial-NoDerivs 2.0 license. Therefore, it can be downloaded without cost from the official website: www.medusabci.com. Additionally, the official repositories for all of our developments, including MEDUSA© Platform and MEDUSA© Kernel, can be accessed at www.github.com/medusabci. Thus, our users have complete access to the source code and can participate in the its development. In our view, the involvement of the community will determine the future success of MEDUSA©. Thus, we put special effort in creating content and tools to promote the participation of the community in the project. Concretely, the website provides updated information, documentation, tutorials, discussion forums and an app market where users can download apps and upload their own projects.

Chapter 5

Results

This chapter presents the most relevant results of this dissertation. It is organized in accordance with the four hypothesis that were outlined in the section 2.1. At the same time, these hypothesis are directly aligned with the research papers included in the compendium of publications (see appendix A). First, SSVEPs elicited by the RCP are analyzed in section 5.1. Second, we assess the performance of the first method proposed to detect the user's control state in ERP-based BCIs in section 5.2. Then, sections 5.3 and 5.4 describe the results achieved by EEG-Inception in the command decoding and control state detection tasks, respectively. Finally, the characteristics and functionalities of MEDUSA©, our novel BCI platform, are detailed in section 5.5.

5.1 SSVEPs elicited by the RCP

In Santamaría-Vázquez et al. (2019a), we studied the effect of the RCP stimulation on the user's EEG. The goal was to find and characterize patterns of brain activity that could later be used to detect the user's control state. In this regard, it was already known that the stimulation frequency of the RCP was somehow reflected in the EEG and could be used to detect the user's control state (Pinegger et al., 2015). However, the underlying mechanisms and characteristics of this phenomenon had not yet been thoroughly investigated. Our hypothesis was that non-target stimuli would elicit a weak SSVEP in the EEG of the user. When a user wants to select a command, they fix their gaze in the target letter or icon. However, non-target stimuli are also perceived through their peripheral field of vision. If the rows and columns of the matrix are highlighted at a fixed rate, these stimuli would trigger

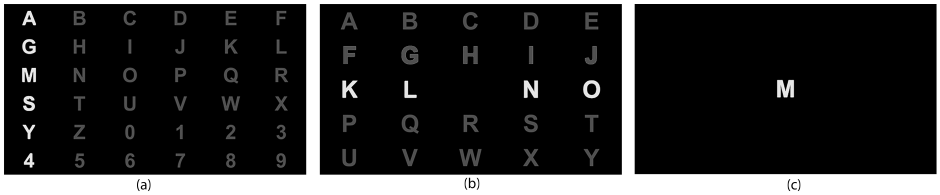

Figure 5.1: Matrices used during the experiments: (a) RCP matrix, (b) overt matrix, (c) covert matrix.

a SSVEP.

To study this phenomenon, we performed a characterization experiment with 5 healthy subjects of the Asynchrony database. This experiment was composed by two analysis aimed to study the origin of the SSVEP and how its characteristics change as a function of the stimulation frequency:

- The first analysis aimed to examine how the characteristics of the SSVEP varied as a function of the stimulation frequency. The stimulation frequency is the inverse of the time interval between two consecutive flashes: $f_{st} = 1/\text{SOA}$. For this analysis, the participants performed six runs of six trials with 15 stimulation sequences using the matrix shown in Figure 5.1(a). The stimulation frequency for each run is listed in Table 5.1.

- The second analysis aimed to investigate the origin of the SSVEP using four runs with different stimulation matrices. Each run consisted of six trials with 15 sequences. The SOA was fixed at 175 ms, corresponding to a stimulation frequency of 5.71 Hz. The first run used the overt matrix shown in Figure 5.1(b), where participants only saw stimuli in their peripheral visual field. The second run used the covert matrix shown in Figure 5.1(c), where participants only saw stimuli in the central region of their visual field. Then, users performed a control and a non-control run with a RCP matrix, depicted in Figure 5.1(a). The study sought to determine whether the SSVEPs are triggered by the peripheral stimuli of the RCP and whether the EEG signal in different modes (overt, covert, control, and non-control) would show the presence or absence of SSVEPs and ERPs with the P300 component.

Figures 5.2 and 5.3 show the results of the first and second analysis of the characterization experiment, respectively. Concretely, the Figure 5.2 characterizes the EEG signal for each stimulation rate. The left graphs depict the average PSD of all trials, channels and participants. The right graphs show the topographic

Table 5.1: Stimulation parameters for the characterization experiment.

Run	SOA	Inter-stimulus time	Stimulus duration	Flashing frequency
1	475 ms	400 ms	75 ms	2.10 Hz
2	375 ms	300 ms	75 ms	2.67 Hz
3	275 ms	200 ms	75 ms	3.64 Hz
4	175 ms	100 ms	75 ms	5.71 Hz
5	150 ms	75 ms	75 ms	6.66 Hz
6	125 ms	50 ms	75 ms	8.00 Hz

SOA: stimulus onset asynchrony

plots of the PSD at the stimulation frequency, averaged for all trials, channels and participants. As hypothesized, the experiment triggered an SSVEP whose main component coincides with the stimulation rate. The harmonics are also visible in the PSD graphs. Interestingly, the frequency and spatial characteristics of the SSVEP changed significantly for the different frequencies.

In Figure 5.3, the brain activity of a single participant in channel Cz is depicted for the four scenarios of the second analysis: overt, covert, control, and non-control. The top of the figure presents the averaged epochs lasting 1000 ms from the start of RCP stimuli. On the other hand, the bottom row displays the average PSD of the trials. As can be seen, the SSVEP is clearly visible in the overt and control conditions, which proves that the origin of this waveform are the non-target stimuli of the RCP perceived through the peripheral field of vision.

The results of this experiment shed light on the origin and characteristics of these weak SSVEPs that are generated when users are controlling an ERP-based speller with the RCP, proving that this side effect can be used to differentiate control vs non-control states (Santamaría-Vázquez et al., 2019a). Furthermore, this control signal is independent of the P300, which is used to decode the commands in these systems and does not allow to accurately differentiate these states due to its low SNR.

5.2 Oddball Steady Response Detection (OSRD) method

Based on the results of the previous experiment, we proposed a novel feature-engineering method to detect the user's control state in ERP-based spellers (Santamaría-Vázquez et al., 2019a). The method, called Oddball Steady Response Detection (OSRD), detects the SSVEPs elicited by non-target stimuli while the user attends to the RCP.

The OSRD method provides a binary output $y \in \{0, 1\}$ after analyzing the

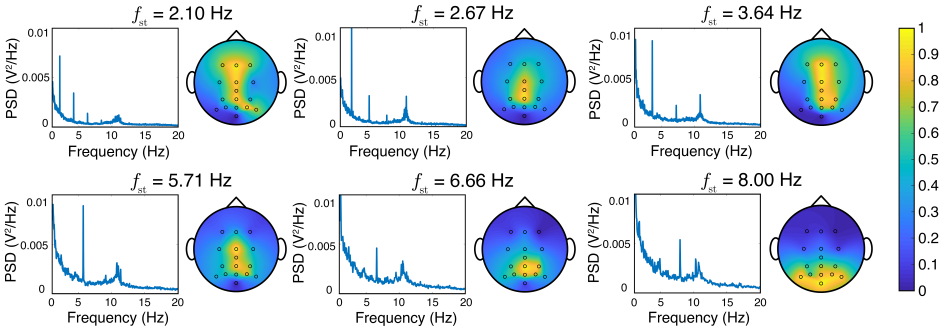

Figure 5.2: SSVEPs for different stimulation frequencies. Graphs on the left show the grand average of the PSD for all trials, channels and participants. Topographic plots on the right show the normalized peak value of the PSD at the stimulation frequency averaged for all trials and participants.

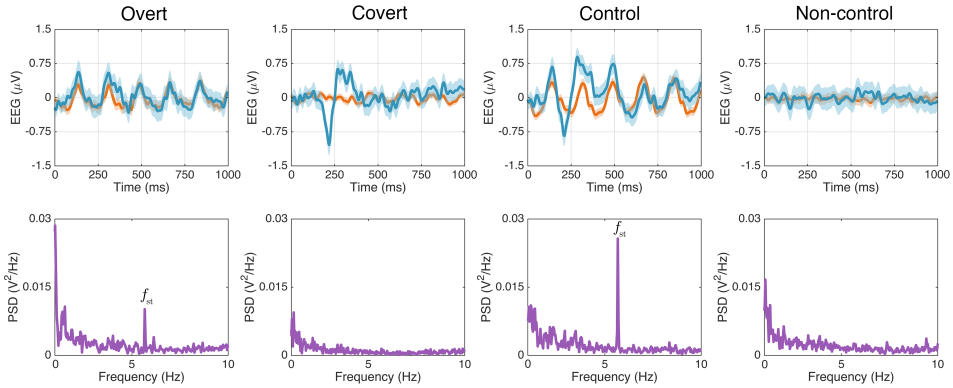

Figure 5.3: Temporal and spectral representation of the ERPs in overt, covert, control and non-control modes for participant S01. The upper part of the figure shows the averaged epochs (i.e., 1000 ms after stimuli) in channel Cz. The shaded area represents the 95% confidence interval. The bottom part of the figure shows the averaged PSD in channel Cz. The stimulation frequency in overt and control modes was 5.71 Hz.

EEG signal corresponding to one trial from the first stimulus of the first sequence to the last stimulus of the last sequence. When the control state is detected ($y = 1$), the system continues to the command decoding pipeline and performs a selection. Conversely, when the non-control state is detected ($y = 0$), no further actions are taken and a new trial begins.

In order to train and test the algorithm, we performed experiments with 15 subjects of the Asynchrony database, as detailed in section 3.2. Regarding the configuration of the system, we used a SD of 75ms and an ISI of 100ms, which gives a SOA of 175ms (see 3.2). Therefore, the corresponding stimulation frequency

of the RCP was $f_{st} = 1/SOA = 5.71$ Hz.

The signal processing methods that compose the OSRD method were explained in section 4, but there are some details that must be further specified. For the pre-processing, OSRD applies bandpass FIR filtering between $[f_{st} - 1,\ f_{st} + 1]$ Hz, where f_{st} is the stimulation frequency (i.e., 5.71 Hz), and CAR spatial filtering . The features that were extracted for each control and non-control trial were the frequency- and correlation-based features that were explained in sections 4.2.2 and 4.2.3, respectively. We proposed two approaches to calculate the features. The first and most classical one used control and non-control trials to calculate the training and testing observations. On the other hand, the second approach used synthetic non-control observations from control trials, avoiding de need of recording non-control trials and thus reducing the calibration time of the OSRD method by half. This approach assumes that the characteristics of the EEG around f'_{st} for control trials are the same as around f_{st} for non-control trials, being $f'_{st} = f_{st} + 0.5$ Hz. In this case, non-control observations can be simulated by calculating the features with control trials, but using f'_{st}. Finally, the features were classified using LDA, as explained in section 4.4.1. This model was trained with control and real non-control/synthetic non-control observations.

In this study, we also applied a classical and widely used processing pipeline for the command decoding task in RCP spellers (Krusienski et al., 2008). This pipeline is independent of the OSRD method. The pre-processing consisted on FIR filtering between 1 and 10 Hz and CAR spatial filtering . The feature-extraction stage used time-based features (see 4.2.1), with sub-sampling to 20 Hz, epochs of 800 ms from the stimulus onset, and baseline correction using the signal 250 ms prior to stimulus onset. The 60 most relevant features were selected using SW as explained in section 4.3.1, and the final ERP classification was performed with LDA. The algorithm selected the command corresponding to the row and column that reached the highest score in the classification stage (Santamaría-Vázquez et al., 2019a).

In sections 5.2.1 and 5.2.2, the results of the experiments that were performed to validate the OSRD method are presented.

5.2.1 Offline experiment

The offline experiment was designed to assess the performance of the OSRD method, validate the synthetic non-control observations approach, and gather data to train the OSRD method for the online experiment.

The 15 participants completed two sessions, each one with five control and five non-control runs of six trials. The control runs required the participants to spell six random characters using a 6×6 RCP matrix, while the non-control runs required them to ignore stimuli while watching a video or reading a text. The experiment setup is depicted in Figure 5.4. Therefore, the dataset of the offline experiment had 60 control trials and 60 non-control trials per participant. Two analyses were conducted to evaluate the OSRD method and validate the synthetic non-control observations approach. The first analysis applied LOO cross-validation (see section 4.5.3) to the entire dataset for each participant (i.e., 119 trials for training, 1 for testing), using both control and non-control trials. The second analysis modified the evaluation procedure by training the OSRD method on synthetic non-control observations, which were created from control trials following the process explained above. Importantly, the testing synthetic non-control observations were replaced by actual non-control observations in each iteration of the LOO process, providing a realistic estimation of the model's performance.

The results of these analyses are summarized in Table 5.2. The test accuracy in the control state detection task using the OSRD method is shown for 1, 5, 10 and 15 sequences, breaking down the results using real and simulated non-control observations for training. As can be seen, the OSRD method achieved a high accuracy, reaching values near 90% with only five sequences of stimulation. Additionally, we analyzed the differences between the real and synthetic non-control observations approaches with Wilcoxon signed-rank test, correcting the FDR with the Benjamini-Hochberg approach, as described in section 4.5.2. We did not found statistical differences, which validates the use of synthetic non-control observations to reduce the calibration time of the asynchronous system.

5.2.2 Online experiment

The online experiment was designed to evaluate the efficacy of OSRD in a more real scenario. To this end, the same participants of the offline experiments performed a third session. In this session, they were asked to make four runs of six selections each (i.e., 24 trials) using the same 6×6 RCP matrix. The participants had to attend stimuli the first three trials of each run, while they had to ignore the stimuli in the others.

For each trial, the control state was determined using the OSRD method. Afterward, the command was selected by applying the command decoding pipeline. Of note, the system only selected a command if control state was determined by

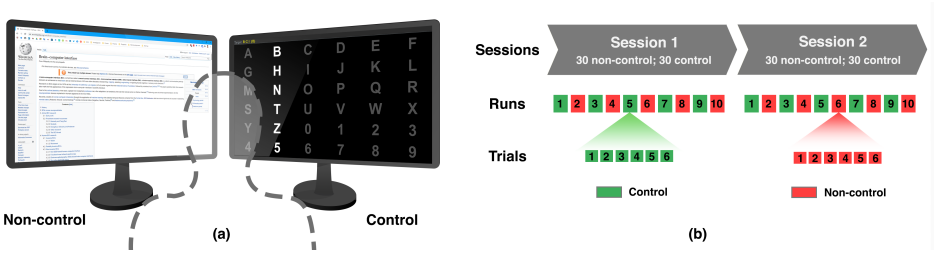

Figure 5.4: Experimental setup. (a) Depicts a schematic representation of the subject and the two screens. The left screen displayed the browser used during the non-control runs while the speller was shown on the right screen. The paradigm was active during both tasks, but subjects were only required to pay attention to the stimuli during the control runs. (b) Shows the experiment plan, which was made up of two sessions, each with 10 runs of 6 trials and 15 sequences. Control and non-control runs were alternated to prevent subject fatigue.

Table 5.2: Results of the offline experiment

	No. Sequences							
	1		**5**		**10**		**15**	
	R	S	R	S	R	S	R	S
U01	55.0	54.2	84.2	85.0	98.3	98.3	99.2	99.2
U02	60.8	59.2	88.3	89.2	95.8	95.8	97.5	97.5
U03	63.3	64.2	94.2	95.0	97.5	97.5	100	100
U04	64.2	64.2	83.3	84.2	96.7	95.8	99.2	99.2
U05	61.7	60.8	87.5	87.5	96.7	96.7	97.5	98.3
U06	65.8	64.2	95.8	95.0	96.7	96.7	98.3	98.3
U07	56.7	58.3	75.0	75.8	79.2	76.7	87.5	86.7
U08	63.3	62.5	97.5	97.5	100	100	100	100
U09	61.7	60.0	89.2	88.3	91.7	92.5	95.0	97.5
U10	50.8	50.0	80.8	80.0	88.3	89.2	88.3	88.3
U11	65.0	65.8	91.7	91.7	100	100	100	100
U12	60.0	61.7	79.2	80.0	86.7	85.8	92.5	92.5
U13	65.0	65.0	79.2	78.3	90.0	90.0	95.0	95.0
U14	60.0	56.7	94.2	95.0	99.2	98.3	100	100
U15	73.3	72.5	96.7	96.7	99.2	99.2	100	100
Mean	**61.8**	**61.3**	**87.8**	**87.9**	**94.4**	**94.2**	**96.7**	**96.8**
±SD	5.0	5.2	7.0	6.9	5.8	6.2	4.1	4.2

Control state detection accuracy (Acc_{csd}). R: training and testing with real non-control observations, S: training with synthetic non-control observations and testing with real non-control observations.

the OSRD method previously. The control state detection and command selection stages were calibrated using the 60 control trials that were collected in the offline experiment for each participant. Therefore, we used synthetic observations to calibrate the OSRD method for the online session. The number of stimulation sequences for each participant was determined as the minimum to reach a training accuracy of 95% in command selection. If a participant could not reach this threshold, the number of sequences was set to 15.

The results of this analysis are summarized in Table 5.3. This table provides a detailed description of the system's performance. Results for the control state

Table 5.3: Results of the online experiment

	N_s	Control State Detection					Overall	
		Acc_{csd}	PPV	NPV	TPR	TNR	Acc_{ovr}	ITR
U01	10	100	100	100	100	100	91.7	12.4
U02	10	95.8	100	92.3	91.7	100	91.7	12.4
U03	8	100	100	100	100	100	95.8	16.8
U04	10	91.7	100	85.7	83.3	100	91.7	12.4
U05	12	91.7	100	85.7	83.3	100	83.3	8.7
U06	10	95.8	100	92.3	91.7	100	91.7	12.4
U07	15	95.8	92.3	100	100	91.7	95.8	9
U08	8	95.8	92.3	100	100	91.7	95.8	16.8
U09	12	95.8	92.3	100	100	91.7	95.8	11.2
U10	15	91.7	91.7	91.7	91.7	91.7	91.7	8.2
U11	10	95.8	100	92.3	91.7	100	95.8	13.4
U12	10	91.7	91.7	91.7	91.7	91.7	91.7	12.4
U13	8	100	100	100	100	100	100	18.5
U14	12	95.8	100	92.3	91.7	100	91.7	10.3
U15	10	95.8	92.3	100	100	91.7	83.3	10.5
Mean	**10.7**	**95.5**	**96.8**	**94.9**	**94.5**	**96.7**	**92.5**	**12.4**
±SD	2.1	2.8	3.9	5.2	5.8	4.1	4.4	2.9

N_s: number of sequences; Acc_{csd}: accuracy of the control state detection stage; PPV: positive predictive value; NPV: negative predictive value; TPR: true positive rate; TNR: true negative rate; ACC_{ovr}: accuracy of the overall system; ITR: information transfer rate (bits/min).

detection stage include Acc_{csd}, PPV, NPV, TPR and TNR achieved by OSRD method during the online sessions. In this analysis, control state has been considered as the positive class and the non-control state as the negative class. Results of the overall system include the control state detection and command detection stages. Hence, control trials were considered correct whether the control state and the command were correctly determined at the same time. For the overall system, ACC_{ovr} and ITR are given. The table shows that the system achieved high accuracy in both tasks, reaching 92.5% for the asynchronous speller.

5.3 Command decoding with EEG-Inception

In Santamaría-Vázquez et al. (2020b), we focused on the ERP detection problem to improve the command decoding accuracy of our asynchronous ERP-based speller. In this study, we proposed a novel deep-learning model, called EEG-Inception, which was described in section 4.4.2.

The model was evaluated in an experiment simulating a realistic scenario, taking into account that the amount of training data from a single subject is limited in real-life settings, and that the target users of ERP-based spellers are severely disabled people. Additionally, we compared EEG-Inception with five previous methods: rLDA (Krusienski et al., 2008), xDAWN + RG (Barachant et al., 2012), CNN-BLSTM (Santamaría-Vázquez et al., 2019b), DeepConvNet (Schirrmeister et al., 2017) and EEGNet (Lawhern et al., 2018).

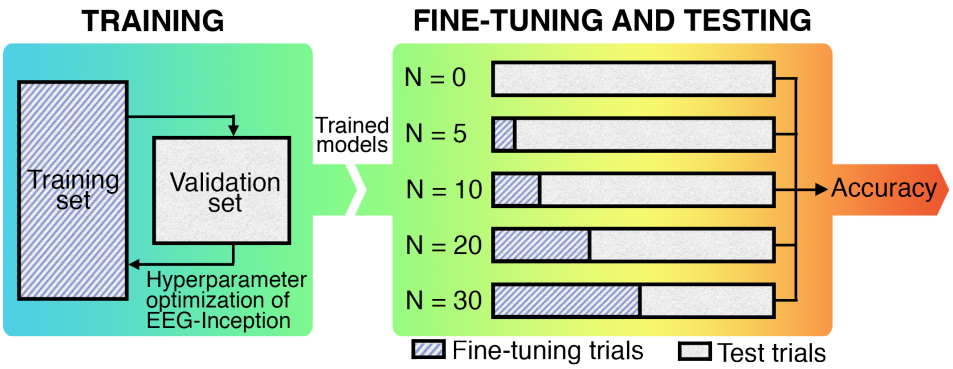

Figure 5.5: Workflow of the study. First, the hyperparameter optimization process for EEG-Inception was done based using the validation set. Then, the models were trained using the training set. Finally, they were evaluated in the test set using a fine-tuning process for each subject. Fine-tuning trials were randomly selected, and the process is repeated 100 times for each N and subject.

For the evaluation experiment, we used the three databases that were detailed in section 3.2. The dataset was split in three. Of the 42 healthy subjects, 33 (i.e., 80%) were assigned to the training set, and 11 to the validation set (i.e., 20%). The remaining 31 severely disabled subjects from the BCI Web Browser database and the BCI Social Networks database were assigned to the test set in order to achieve a more realistic estimation of the performance of the model in this population.

The hyperparameter optimization process and the evaluation experiment are detailed in sections 5.3.1 and 5.3.2. The Figure 5.5 shows a schematic representation of the process followed to perform these analyses.

5.3.1 Hyperparameter optimization

We optimized the learning rate (lr), activation function (f_{act}) and dropout rate (p_{drop}) of EEG-Inception using the validation set. The search space for each hyperparameter was: $lr = \{0.01, 0.001, 0.0001\}$; $f_{\text{act}} = \{\text{Sigmoid}, \text{ReLU}, \text{ELU}\}$; $p_{\text{drop}} = \{0.00 : 0.05 : 0.5\}$. The Figure 5.6 depicts the results of this optimization process. The score is the command decoding accuracy in the validation set. The optimal values for each hyperparameter were chosen by identifying the set that resulted in the highest score.

The optimal set was $lr = 0.001$; $f_{\text{act}} = ELU$; $p_{\text{drop}} = 0.25$. The results also showed that the activation function and learning rate had great impact in the performance of EEG-Inception, whereas dropout rates between 0.1 and 0.4 yielded

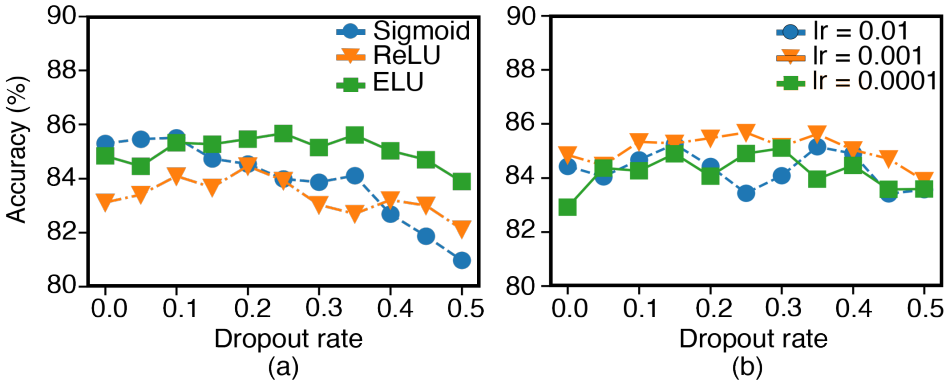

Figure 5.6: Results of the optimization process of EEG-Inception in the validation set (8 healthy subjects). Each line depicts the mean command decoding accuracy after a fine-tuning process for each subject with $N = 30$, considering 5 sequences of stimulation. (a) Score for each activation function and dropout rate using the best value of learning rate ($lr = 0.001$). (b) Score for each learning rate and dropout rate using the best activation function (ELU).

similar results.

5.3.2 Evaluation experiment

The performance of EEG-Inception and the comparison models was assessed with the evaluation experiment. Firstly, all models were trained using the training set. Then, we evaluated all the models in the test set using a fine-tuning process for each subject with $N = \{0, 5, 10, 20, 30\}$ trials. When $N = 0$, the models were directly evaluated on the test set, simulating a plug & play device. This allows us to assess their robustness to inter-subject variability. For $N > 0$, the models were fine-tuned. First, the algorithm picked N trials from each test subject randomly. Then, the model, which is already initialized with the training set, fits the data from these trials. Finally, the fine-tuned model is tested with the rest of trials of the test subject. This process was repeated 100 times for each value of N and subject, and the results were averaged to obtain a final performance measure.

The evaluation experiment was designed to mimic a real-life scenario, where the calibration time should be under 20 minutes of effective training (or approximately 30 trials) to maintain a suitable usability. In this regard, excessively long calibration times can negatively impact the usability of an ERP-based speller (Martínez-Cagigal et al., 2017). Our training approach, which combines cross-subject transfer learning with fine-tuning, aims to reduce the required calibration

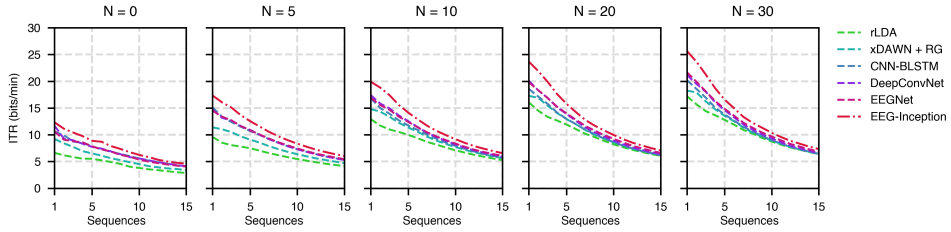

Figure 5.7: ITR: information transfer rate; N: number of fine-tuning trials. ITR in bits/min in the test set (31 motor disabled subjects) as a function of N and the number of sequences.

time and prevent overfitting. Moreover, the distribution of the dataset for the evaluation experiment, with all the severely disabled subjects in the test set, allowed us to evaluate the effectiveness of the proposed model and training strategy for end users of ERP-based spellers.

Table 5.4 summarizes the results of the evaluation experiment, showing the command decoding accuracy of the models in the test set for each number of fine-tuning trials and sequences considered to make the prediction. The last column provides the improvement achieved by EEG-Inception with respect to the rest of models. This improvement was calculated as the mean difference in the command decoding accuracy. Wilcoxon Signed-Rank Test was used to assess the statistical difference between the performance of EEG-Inception and the other models, correcting the FDR with the Benjamini-Hochberg approach, as detailed in 4.5.2. On the other hand, the Figure 5.7 shows the simulated ITR of the models.

As can be seen, EEG-Inception clearly outperformed the rest of models, reaching the highest performance for command decoding accuracy and ITR. Moreover, this improvement was significant (p-value < 0.01), regardless of the number of fine-tuning trials and sequences considered in the analysis.

5.4 Control state detection with EEG-Inception

In Santamaría-Vázquez et al. (2020b), we focused on increasing the command decoding accuracy of our ERP-based speller with EEG-Inception. However, we did not address the control state detection problem. EEG-Inception could be complemented with the OSRD method to provide an asynchronous control. However, after the impressive results of this CNN in the ERP detection task, our goal was to achieve similar improvements in the control state detection stage. Consequently, we continued with this line of research in Santamaría-Vázquez et al. (2022). Unlike

Table 5.4: Command decoding accuracy

N	Model	No. Sequences 1	5	10	15	Imp.
0	rLDA	14.6 ± 8.6	33.1 ± 20.4	40.3 ± 24.2	43.7 ± 25.1	**14.7**
	xDAWN + RG	17.9 ± 9.2	37.5 ± 19.6	46.2 ± 22.8	49.7 ± 23.4	**9.6**
	CNN-BLSTM	19.6 ± 8.8	42.7 ± 19.8	53.6 ± 22.0	57.2 ± 22.5	**3.7**
	DeepConvNet	20.7 ± 9.8	42.0 ± 20.8	53.2 ± 23.4	56.9 ± 23.7	**4.0**
	EEGNet	19.5 ± 9.5	42.3 ± 21.0	51.9 ± 23.2	56.5 ± 23.3	**4.6**
	EEG-Inception	21.3 ± 9.9	46.1 ± 21.1	57.4 ± 23.7	61.4 ± 23.8	—
5	rLDA	18.6 ± 8.6	41.7 ± 18.7	53.4 ± 20.3	58.2 ± 20.0	**16.0**
	xDAWN + RG	20.6 ± 8.9	47.3 ± 19.6	59.0 ± 20.5	63.4 ± 19.9	**10.7**
	CNN-BLSTM	24.3 ± 10.3	52.7 ± 19.8	64.7 ± 20.2	68.5 ± 19.4	**5.6**
	DeepConvNet	23.8 ± 10.3	53.2 ± 20.5	64.3 ± 21.3	68.3 ± 20.4	**5.7**
	EEGNet	23.8 ± 9.8	53.0 ± 20.2	65.3 ± 20.8	69.4 ± 20.1	**5.1**
	EEG-Inception	26.4 ± 10.8	58.7 ± 19.4	70.7 ± 18.8	74.6 ± 17.2	—
10	rLDA	22.0 ± 10.0	50.4 ± 19.5	63.5 ± 20.3	68.2 ± 19.2	**11.6**
	xDAWN + RG	24.1 ± 9.8	54.9 ± 19.4	67.2 ± 19.1	71.5 ± 17.8	**7.7**
	CNN-BLSTM	25.9 ± 10.7	56.0 ± 19.9	68.3 ± 19.5	72.1 ± 18.5	**6.7**
	DeepConvNet	26.4 ± 11.0	58.4 ± 20.3	69.6 ± 20.1	73.3 ± 18.8	**5.2**
	EEGNet	26.1 ± 10.2	58.2 ± 19.8	70.6 ± 19.6	74.5 ± 18.3	**4.8**
	EEG-Inception	28.8 ± 11.1	63.6 ± 18.7	75.3 ± 17.6	78.9 ± 15.5	—
20	rLDA	25.1 ± 11.0	56.7 ± 19.6	69.9 ± 19.6	74.6 ± 17.8	**9.9**
	xDAWN + RG	26.4 ± 10.4	60.1 ± 18.3	72.3 ± 17.3	76.4 ± 15.6	**7.2**
	CNN-BLSTM	27.5 ± 11.2	60.1 ± 19.2	72.3 ± 18.1	75.7 ± 16.8	**7.2**
	DeepConvNet	28.9 ± 11.6	62.7 ± 19.6	73.7 ± 18.7	77.1 ± 17.0	**5.6**
	EEGNet	28.8 ± 10.7	63.2 ± 18.8	75.6 ± 17.9	79.1 ± 16.2	**4.3**
	EEG-Inception	32.0 ± 12.1	68.4 ± 17.9	79.4 ± 16.0	82.8 ± 13.7	—
30	rLDA	26.2 ± 11.2	59.5 ± 19.3	72.7 ± 19.0	77.3 ± 16.9	**9.4**
	xDAWN + RG	27.3 ± 10.6	62.0 ± 17.9	74.2 ± 16.5	78.2 ± 14.9	**7.5**
	CNN-BLSTM	29.0 ± 11.5	62.6 ± 18.9	74.4 ± 17.4	77.7 ± 16.0	**7.2**
	DeepConvNet	29.8 ± 11.8	64.9 ± 19.1	75.8 ± 17.9	79.0 ± 16.3	**5.7**
	EEGNet	30.4 ± 11.3	65.7 ± 18.6	77.8 ± 17.2	81.0 ± 15.5	**4.3**
	EEG-Inception	33.7 ± 12.4	70.6 ± 17.4	81.5 ± 15.3	84.6 ± 13.2	—

N: number of fine-tuning trials for each subject. Command decoding accuracy (Acc_{cmd}) averaged over the test set subjects (31 motor disabled). Column "Imp." shows the accuracy improvement (%) yielded by EEG-Inception compared to the other 5 models, calculated as the mean difference in command decoding accuracy for each N. Statistical differences between EEG-Inception and the other models were assessed with Wilcoxon Signed Rank Test, correcting the FDR with Benjamini-Hochberg approach, as described in section 4.5.2. All comparisons were significant (p-value < 0.01), regardless of N and the number of sequences.

earlier deep-learning approaches for ERP classification, EEG-Inception allows the analysis of EEG signals in multiple temporal scales (Santamaría-Vázquez et al., 2020b). This could enhance the model's versatility, making it more adaptable to a wider range of tasks. Therefore, our hypothesis was that this architecture could be leveraged to provide a reliable asynchronous control of ERP-based spellers, targeting different EEG patterns linked to the control of these systems, such as SSVEPs elicited by the RCP (Santamaría-Vázquez et al., 2019a), or differences in the complexity of the EEG due to the user's concentration, especially in the prefrontal cortex (Martínez-Cagigal et al., 2019b). In this regard, neither EEG-

Inception nor any other deep-learning model had yet been used to discriminate the user's control state in ERP-based spellers.

To test this hypothesis, we performed an experiment designed to evaluate the performance of EEG-Inception for this task. First, we extended the Asynchrony database, which had 15 subjects at that moment, to reach 22 subjects, following the same protocol as in the OSRD study. This protocol is described in section 5.2.1. Then, we adapted the signal processing pipeline presented in Santamaría-Vázquez et al. (2020b) to facilitate the integration of EEG-Inception for this control state detection. This pipeline used the same features for the control state detection and command decoding tasks. The pre-processing stage applied FIR filtering between 0.5 and 45 Hz to keep as much discriminative information as possible. Then, a CAR spatial filtering was used to remove noisy artifacts. The features were calculated as described in 4.2.1. We applied decimation to 128 Hz and extracted the epochs from 0 to 1000 ms after the onset. Baseline correction was performed taking a window of 250 ms before the stimulus. Thus, the feature array had 128 samples $\times$ 8 channels, which are the input dimensions required by EEG-Inception (Santamaría-Vázquez et al., 2020b). Considering that the experiment had 2 sessions of 10 runs, with 6 trials of 15 sequences per run, the number of observations for each subject was 21,600. Then, the total number of EEG epochs for the 22 subjects was 475,200. Finally, two different EEG-Inception instances are used for each of the two classification tasks.

The control state detection stage dynamically turns the system into an asynchronous BCI. The workflow in this stage for each trial was as follows: (1) the epochs are classified with the model trained to discriminate the control state; (2) the model returns the probability of each state for every observation; and (3) the probabilities are averaged, determining non-control state when the score is less than 0.5 and control state otherwise. If the trial is classified as non-control, the system starts a new trial without taking further actions. On the other hand, if the trial is classified as control, the system continues to the command decoding stage. Note that the algorithm operates under the assumption that all observations within a given trial have the same control state. Consequently, users are advised to avoid switching tasks until the current trial has been completed. This will ensure that the proposed method can analyze the data accurately and provide reliable results.

Once the control state was determined for a trial, the other instance of EEG-Inception is used to decode the command. This model was trained to detect P300 potentials. Thus, it discriminates between target and non-target epochs of the

RCP. As the output scores of the model are associated to the row or column of the stimulus, the system selects the command with maximum probability, providing the appropriate feedback to the user. This process marks the beginning of a new trial, and the cycle is repeated.

5.4.1 Evaluation experiment

In order to train and validate our approach, we simulated the real use of the speller using LOO cross-validation combined with cross-subject transfer learning and fine-tuning. The evaluation process was similar to the one used to validate EEG-Inception in Santamaría-Vázquez et al. (2020b). For each iteration, the training subjects were used to initialize the models of the control state detection and command decoding stages. This implies that the model for control state detection was trained with 453,600 observations (21 subjects $\times$ 120 trials $\times$ 180 observations/trial), whereas the model for command decoding was trained with 226,800 observations (21 subjects $\times$ 60 control trials $\times$ 180 observations/trial). Then, $N = \{0, 5, 10, 20, 30\}$ control and non-control trials from the test subject were randomly selected to fine-tune the models. Of note, these trials were not used in the test phase. As a result, the number of test trials for each N was the difference between the total number of trials and the number of fine-tuning trials. To obtain a reliable validation, this process was repeated 100 times for each subject, averaging results. By means of this analysis, it is possible to study how the performance of the system is affected by the number of fine-tuning trials. $N = 0$ represents a plug-and-play device with no calibration for the user, whereas $N = 30$ would require a 30-minute calibration session to perform the control and non-control selections.

Regarding the training process, the two models were trained independently using different labels. Specifically, for the control state detection model, EEG epochs were categorized as either control or non-control. The epochs in which the user was paying attention to the stimuli were classified as control (positive class), while those in which the user was focused on the web browser were classified as non-control (negative class), thus generating a balanced dataset. Regarding the command decoding model, it was trained with control trials only. In this case, epochs corresponding to target commands were classified as P300 (positive class), while non-target epochs were classified as non-P300 (negative class). Both models were trained using identical settings: Adam optimizer with $\beta_1 = 0.9$ and $\beta_2 = 0.999$; categorical cross-entropy loss; batch size of 1024; and a maximum of

Table 5.5: Control state detection accuracy using EEG-Inception

N	No. Sequences														
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
0	81.4	85.23	86.63	86.89	87.01	87.69	88.26	88.22	87.95	88.48	89.02	89.05	89.28	89.13	89.36
5	86.11	88.96	89.99	90.62	90.98	91.06	91.29	91.45	91.63	91.74	91.89	91.91	91.98	92	92.05
10	88.33	91.16	92.27	92.84	93.23	93.47	93.61	93.82	94.05	94.14	94.23	94.3	94.29	94.42	94.48
20	91.24	94.37	95.31	95.76	96.06	96.31	96.45	96.6	96.69	96.83	96.98	97.06	97.07	97.21	97.28
30	91.91	94.41	95.23	95.67	96.02	96.33	96.53	96.53	96.52	96.66	96.77	96.76	96.85	96.88	96.95

N: number of fine-tuning trials in control state for each subject. Thus, the total number of calibration trials used to fine-tune the model for this task was $2N$ (i.e., N control, N non-control). Test accuracy for the control state detection task (Acc_{csd}) averaged over the 22 subjects.

500 training iterations over the entire dataset, applying early stopping if there was no improvement for 10 consecutive iterations (Santamaría-Vázquez et al., 2020b).

The results of the control state detection task and the overall asynchronous system are presented separately. The Table 5.5 and the Figure 5.8 summarizes the results of the cross validation experiment for the control state detection task. Concretely, the Figure 5.8 shows the normalized confusion matrices, whereas the Table 5.5 shows the accuracy averaged across all subjects and broken down by the number of fine-tuning trials and stimulation sequences. Together, they give a complete overview of the system's performance in this task. Additionally, Figure 5.9 characterizes the EEG in time and frequency to analyze differences between correctly and incorrectly classified trials in the control state detection task and understand which could be the main factors affecting the performance of the model. As can be seen, the proposed method was able to discriminate the control state with high accuracy.

Regarding the results of the overall system, the Table 5.6 shows the accuracy including both stages for control state detection and command decoding, considering that both classifications must be correct at the same time. Therefore, if one stage fails, it is considered as a mistake. As before, the accuracy is averaged across subjects and broken down by the number of fine-tuning trials and stimulation sequences. Additionally, Figure 5.10 shows the simulated ITR that would be achieved by the asynchronous speller in an online experiment. This metric takes into account the speed and the accuracy of the system, allowing a direct comparison between different BCIs (Wolpaw et al., 2002). It should be noted that the ITR was calculated only for control trials but considering the overall accuracy, including control state detection and command decoding stages. As before, these analyses show that our approach reached high performance in terms of accuracy and simulated ITR for the overall system.

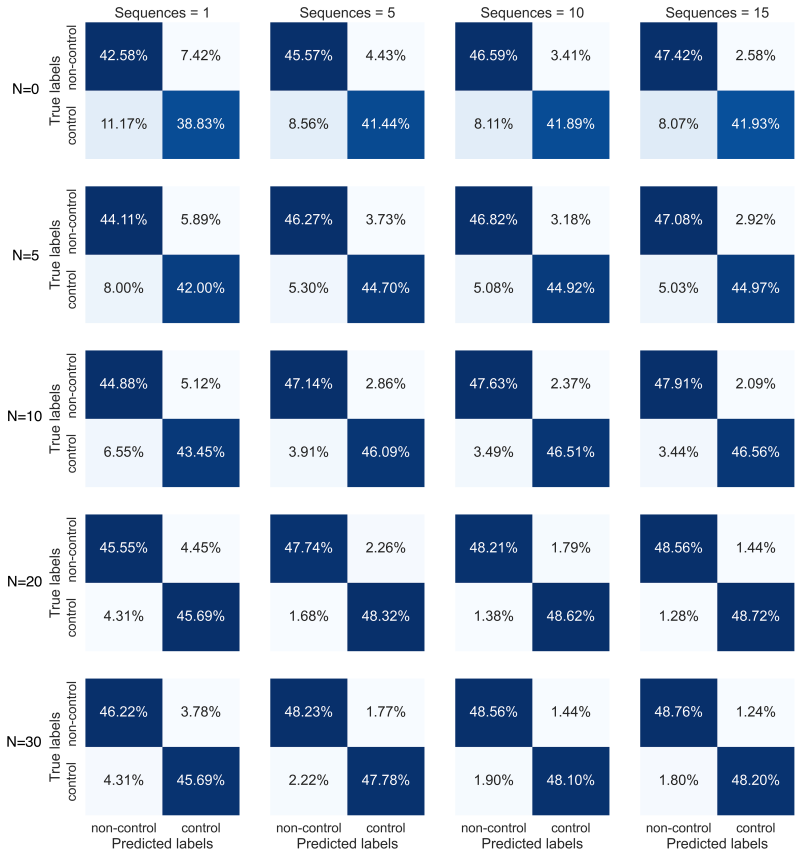

Figure 5.8: Normalized confusion matrices averaged across subjects.

Table 5.6: Overall accuracy of the asynchronous system using EEG-Inception

N	No. Sequences														
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
0	52.05	60.49	66.7	70.00	73.45	75.49	77.77	79.43	79.62	80.95	82.20	83.26	83.94	84.24	85.38
5	59.70	70.30	76.05	80.77	83.48	85.11	86.47	87.76	88.6	89.22	89.72	90.09	90.33	90.68	90.99
10	62.63	73.73	80.34	84.55	87.09	88.75	90.07	91.27	91.99	92.44	92.78	93.11	93.20	93.58	93.79
20	64.90	77.34	83.98	88.20	90.73	92.32	93.32	94.42	94.98	95.49	95.82	96.21	96.20	96.51	96.68
30	66.49	78.20	84.33	88.69	91.30	92.96	93.71	94.61	94.95	95.34	95.70	96.01	96.09	96.30	96.47

N: number of fine-tuning trials in control state for each subject. Thus, the total number of calibration trials used to fine-tune the model for control state detection was $2N$ (i.e., N control, N non-control), whereas the model for command decoding only used N control trials. Overall test accuracy (Acc_{ovr}) in percentage for the control state detection and command decoding tasks averaged over the 22 subjects. Noteworthy, one trial is considered correct only if both conditions were correctly classified at the same time.

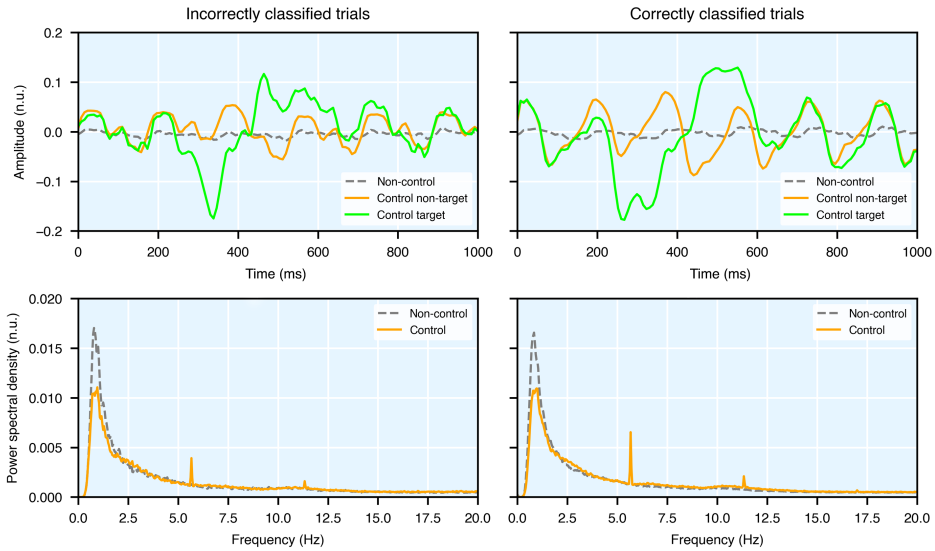

Figure 5.9: n.u: normalized units. Characterization of correctly and incorrectly classified trials for the control state detection task. The upper graphs show the averaged EEG epochs for the 3 different conditions: non-control, control non-target and control target. The lower graphs show the power spectral density of the entire trials.

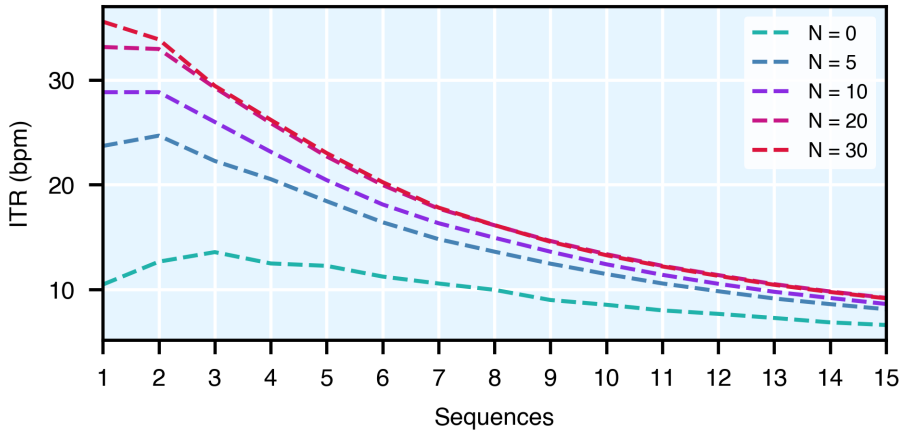

Figure 5.10: ITR: information transfer rate (bpm); N: number of fine-tuning trials in control state for each subject. Average ITR of the overall system, including control state detection and command decoding stages and only considering control trials.

5.5 MEDUSA©: our novel BCI platform

When we were defining the objectives of this dissertation, and after my experience with the bachelor's and master's theses, we realized that the software tools that were available for BCI research were not fully adapted to our needs. As explained in section 1.4.3, we identified a number of drawbacks in the most widespread BCI platforms (e.g., BCI2000, OpenVibe) that could be addressed to facilitate the research in this field. Therefore, we decided to build a novel BCI platform: MEDUSA©. At the beginning, our goal was to implement the necessary functions to develop and test the proposed ERP-based speller. This first version of the software was presented at a national conference in 2018 (Santamaría-Vázquez et al., 2018). However, as we moved forward in the investigation and we developed new projects, our purpose became increasingly ambitious. We wanted to include new BCI paradigms and cognitive neuroscience experiments and share them with the community to increase the impact of our work. Furthermore, if the community could contribute with their developments too, MEDUSA© could become a new space to share the latest advances in the field. With this vision guiding us, we consistently worked on MEDUSA© throughout the development of this dissertation following the architecture and methodologies that were detailed in section 4.6. The results of these efforts were presented in Santamaría-Vázquez et al. (2023) and are summarized in the following sections.

5.5.1 MEDUSA© Kernel

MEDUSA© Kernel is a Python library, available in the Python Package Index (PyPI) repository, with a complete suite of methods for signal processing. These functions can be categorized according to their different levels of abstraction. The first level is composed of low-level functions, which are generic methods that can be used to process signals in many scenarios, being the most relevant:

- Frequency filters: configurable IIR and FIR filters.
- Spatial filters: CAR, Laplacian filter, multi-class common spatial patterns (CSP) and CCA.
- Local activation metrics: including spectral metrics, such as band power, median frequency or Shannon entropy; and non-linear features, such as central tendency measure, sample entropy, multiscale entropy, Lempel-Ziv's complexity and Multiscale Lempel-Ziv's complexity.
- Connectivity metrics: including amplitude-based metrics, such as amplitude

correlation envelope (AEC) and instantaneous amplitude correlation (IAC); and phase metrics, such as phase locking value (PLV), phase-based lag index (PLI) and weighted PLI (wPLI).

In a higher level of abstraction, there are functions that apply a whole processing pipeline to the input data to analyze certain features. MEDUSA© Kernel does not assume the nature of the input data in low-level functions, but most of the high-level analysis that are currently implemented are designed to work with EEG recordings. In short, high-level functions use the low-level methods to implement specific use-cases (e.g., command decoding task in an ERP-based speller). The current version of MEDUSA© Kernel includes the following modules:

- ERP-based paradigms: the complete classification pipelines for the control state detection and command decoding tasks that were developed in the previous studies, including the OSRD method (Santamaría-Vázquez et al., 2019a), regularized linear discriminant analysis (rLDA) (Santamaría-Vázquez et al., 2020b), EEGNet (Lawhern et al., 2018) and EEG-Inception (Santamaría-Vázquez et al., 2020b), which can be applied in offline and online ERP-based spellers; P300 analysis and charts; and specialized data structures to handle data ERP data.
- MI paradigms: complete classification pipelines including CSP + rLDA (Fu et al., 2019), EEGNet (Lawhern et al., 2018), EEG-Inception (Pérez-Velasco et al., 2022) and EEGSym (Pérez-Velasco et al., 2022) that can be applied in offline and online modes; MI analysis charts; and specialized data structures to handle MI data.
- c-VEP-based paradigms: offline and online circular-shifting reference pipeline based on CCA (Martínez-Cagigal et al., 2021); c-VEP analysis and charts; raster latencies correction; filter banks; and maximal length sequences (i.e., m-sequences) generation through linear feedback shift registers (LSFR) for binary and p-ary bases.
- Neurofeedback: battery of high-level models based on spectral and connectivity metrics ready to be applied in online and offline apps (Marcos-Martínez et al., 2021).

This organization favors the independence between the different components of the library, increasing the modularity. Additionally, some of the functions are designed to be applied in both online and offline experiments. Therefore, this package can be used for offline analysis of previously recorded data, such as public databases, or real-time tasks. This is an interesting feature that allows to

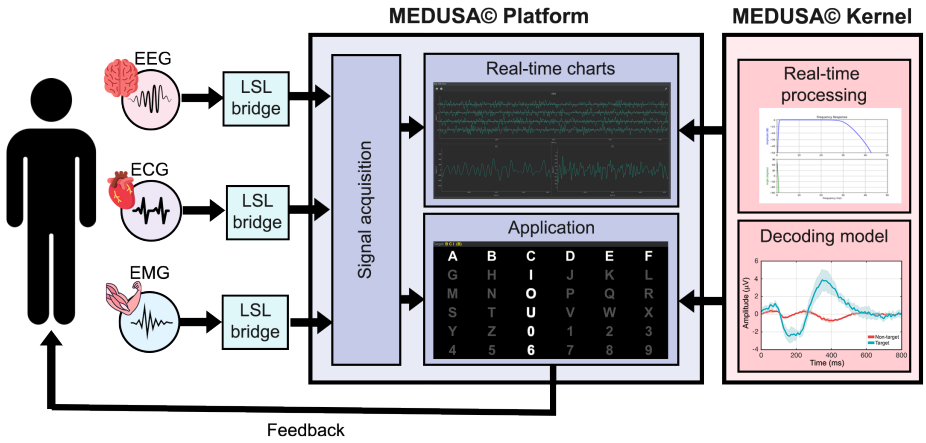

Figure 5.11: Schematic overview of MEDUSA©. EEG: electroencephalography; ECG: electrocardiography; EMG: electromyography; LSL: lab-streaming layer. An arbitrary number of input signals can be received in the platform through the LSL. These signals are available for real-time charts and apps, which implement open-loop or closed-loop BCI and neuroscience experiments. In this case, we represent the row-column paradigm (RCP) app. Functions from the Kernel may be used for signal processing in real-time charts and apps. In this example, the model detects the event-related potentials (ERP) in the EEG.

reproduce exactly the same results in online and offline experiments, which facilitates the reproducibility of a research experiment. Finally, it should be noted that community contributions are welcome. For more information, check the official documentation page at https://docs.medusabci.com/kernel, where the API reference, hands-on tutorials, and the contributor's guide can be found.

5.5.2 MEDUSA© Platform

MEDUSA© Platform has a flexible and adaptable architecture that can be used for a wide range of BCI and neuroscience experiments. Its architecture has three modules: signal acquisition, real-time charts, and apps. The functionalities of these modules can be controlled through a user-friendly GUI that provides a quick and intuitive control of the software. A schematic overview of all the components of MEDUSA© is shown in Figure 5.11. Additionally, the main window of MEDUSA© Platform is shown in Figure 5.12. The official documentation can be found at https://docs.medusabci.com/platform, where some hands-on tutorials, and the contributor's guide can be found.

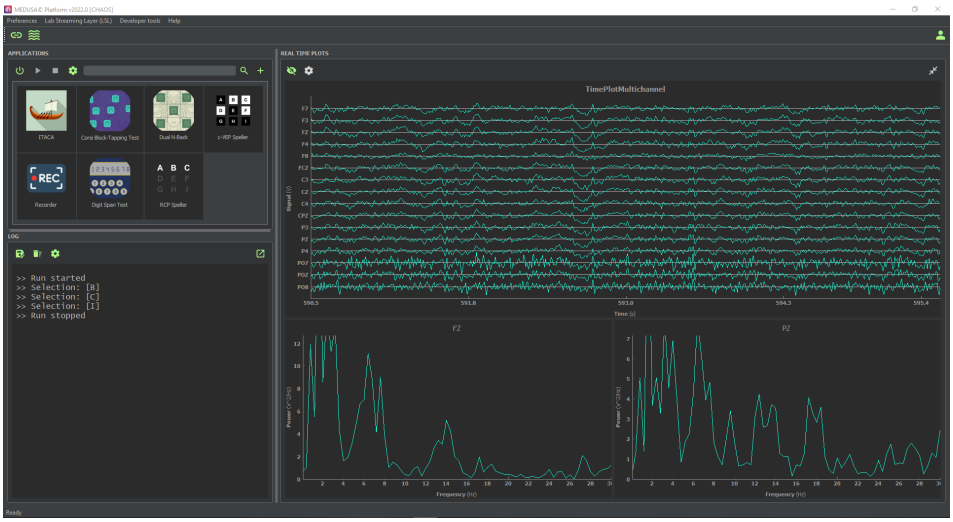

Figure 5.12: Main window of MEDUSA© Platform. The view is divided in panels to control the different functionalities. These panels are: apps (up-left), log (bottom-left) and real-time plots (right). Additionally, more controls and configurations are available in the task bars.

Signal acquisition

MEDUSA© Platform manages the incoming data using the lab-streaming layer (LSL) protocol (Kothe, 2014). This open-source protocol handles the networking, time-synchronization and real-time propagation to allow the transmission of time series data, making our solution independent of the data recording hardware. Additionally, the signal acquisition module wraps the LSL protocol with high-level functionalities to offer advanced data acquisition features.

In order to connect a device with MEDUSA© Platform, an LSL bridge (i.e., independent program or script) is needed to receive data using the specific interface of the device and send it through LSL. Then, MEDUSA© Platform can be configured to receive the stream. Once the stream has been added to the platform's workspace, it is available for the other modules. An important detail is that the signal acquisition module is able to manage several LSL streams at the same time, which is useful in a wide variety of experiments (e.g., synchronous recording of ECG and EEG, implementation collaborative/competitive BCIs).

The LSL protocol is widely used in many research laboratories around the world to record biological signals, having a large community of users and developers. Thus, most commercial EEG recording devices (e.g., g.tec, Brain Products, Neuroscan) have ready-to-use LSL bridges compatible with MEDUSA© Platform.

If a device does not have this support yet, it is possible to create an LSL bridge to connect the device to LSL. These features make LSL an ideal technology for the signal acquisition module of MEDUSA© Platform, allowing it to offer advanced functionality compared to other available solutions.

Real-time charts

MEDUSA© Platform includes time and frequency of charts for visualizing LSL streams in real time using PyQtGraph (Luke Campagnola, 2020), an open source Python package optimized for real-time representations. The panel can be customized to fit various charts that can be adapted to different situations.

Temporal charts display the time courses of signals in real time. Both single and multi-channel charts are supported, with options for pre-processing the signal before representation, adjusting the display time, reducing computational cost through decimation, scaling the display, and customizing the graphical user interface (e.g., line width, colors, etc.).

Frequency charts estimate the PSD of the signal using the Welch method in real time. This lightweight algorithm allows for smooth, real-time visualization of single or multi-channel signals. The charts can be configured with various parameters, such as spectral resolution, overlap, and segment length. These charts provide an alternative way to visualize data and can be used to identify specific features in the signal, noisy components, or sensor malfunctions, among others.

Apps

Apps are the most crucial components of MEDUSA© Platform. These programs carry out tasks or stimulation paradigms, while simultaneously providing real-time feedback and monitoring one or more signals. The platform's architecture has been structured to allow independence between the apps and the other modules. This design enables the development of new apps as needed, without requiring changes to the base code of the platform, increasing the scalability of out solution. Additionally, the software offers generic workflows that can be used in most BCI and neuroscience experiments, which facilitates the implementation of new apps MEDUSA©.

MEDUSA© platform supports apps based on Qt and Unity to handle graphical part of the app (e.g., visual stimuli and feedback presentation). Qt is a C++ library with official Python bindings that is commonly used for developing GUI applications. It offers a wide range of predefined widgets and functions, making it

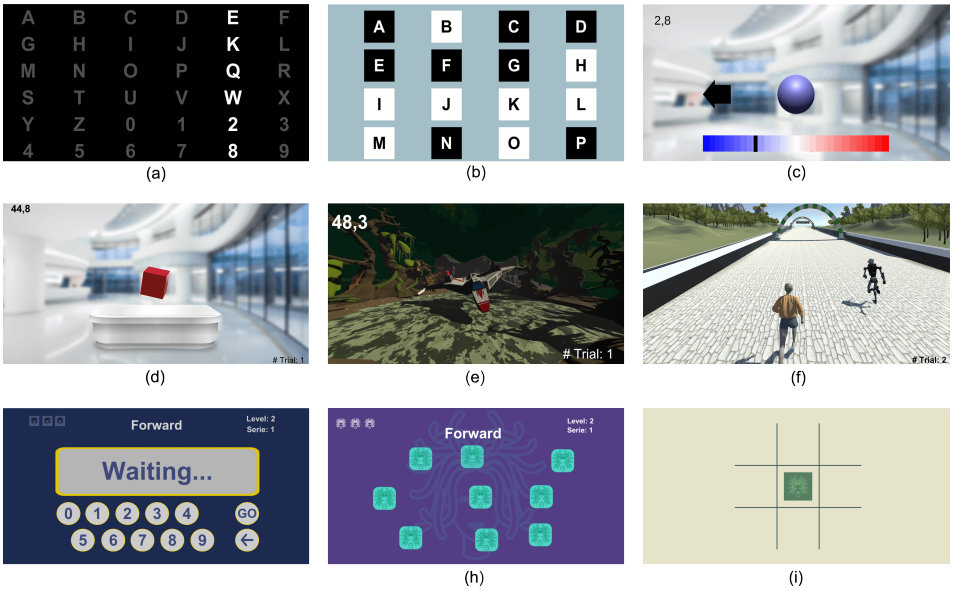

Figure 5.13: Screenshots of some MEDUSA© apps. (a) row-column paradigm (RCP) speller, (b) code-modulated visual evoked potentials (c-VEP) speller, (c) Motor imagery paradigm, (d) Jedi cube scenario from the neurofeedback app, (e) Luke's spaceship scenario from the neurofeedback app, (f) Neurorunner scenario from the neurofeedback app, (g) Digit Span test from the Neuropsychological evaluation app (h) Corsi Block-Tapping test from the Neuropsychological evaluation app, (i) Dual N-Back test from the Neuropsychological evaluation app.

suitable for quickly creating simple apps. However, it does not offer precise refresh rates and time synchronization, which can be important for certain experiments (e.g., c-VEP or SSVEP spellers). On the other hand, Unity is a powerful graphics engine with advanced options for GUI control, 3D modeling, and animations. It requires coding in C#, its native programming language, and is more complex to use for visual applications and games. However, it does provide precise control over refresh rates, making it suitable for advanced BCI applications. The communication between Unity and MEDUSA© is handled through a multi-client asynchronous TCP/IP protocol.

Currently, MEDUSA© platform offers six ready-to-use apps:

- **Recorder.** This app is designed to simplify the process of recording biosignals by providing a generic open-loop system. It allows users to define custom conditions (e.g., eyes closed, eyes open) and events (e.g., movement, blink) and create a recording plan to automatize experiments.

- **RCP speller.** The RCP speller app implements the ERP-based BCI that has been developed in this dissertation. First, it allows to configure all the important parameters: stimulus duration, inter-stimulus interval, text or icon commands, flashing colors or command functions. The app also allows to configure the number of commands, including nested matrices to design complex menus for practical applications. Regarding the signal processing methods, all the models that have been presented in this document are included. For the command decoding task, the available models are the pipelines based on rLDA (Santamaría-Vázquez et al., 2020b), EEGNet and EEG-Inception (Santamaría-Vázquez et al., 2020b). For the control state detection task, we included OSRD (Santamaría-Vázquez et al., 2019a) and the pipeline based on EEG-Inception (Santamaría-Vázquez et al., 2022). Figure 5.13(a) shows a screenshot of this application.

- **c-VEP speller.** The c-VEP speller app is, to the best of our knowledge, the only available open-source implementation of this BCI paradigm. It allows the customization of the number of commands, sequence stimulation rate, sequence length, flashing colors and signal processing methods. It also includes different binary m-sequences with different lengths which are automatically encoded using circular-shifting. The default command decoding pipeline uses a filter bank, epoch averaging and CCA to compute command templates (Martínez-Cagigal et al., 2021). Figure 5.13(b) shows a screenshot of this app.

- **Motor imagery.** The MI app presents feedback to the user (e.g., a colored sphere and a sliding bar) depending on the detected class. It allows to configure different parameters, such as preparation, trial or rest times. The app includes two classification pipelines: CSP + rLDA (Fu et al., 2019) and EEGSym (Pérez-Velasco et al., 2022). The latter model includes an initialized version of EEGSym trained with data from 280 subjects that allows to start controlling the app without prior calibration with an expected accuracy above 80% (more information in (Pérez-Velasco et al., 2022)). The MI application also implements a tool to visualize the ERD/ERS events associated to MI. Figure 5.13(c) shows a screenshot of this app.

- **Neurofeedback.** This app includes three different scenarios to enable the design of progressively more challenging NF protocols. Each scenario has a unique, gamified design to keep users motivated and engaged throughout the study (Roc et al., 2021). It also allows users to easily set up all the key

parameters of a NF study, including the feedback rate, the brain activity to be trained (using metrics based on band powers or connectivity between different regions), and the difficulty of the training. Figures 5.13(d), 5.13(e) and 5.13(f) show screenshots of the three training scenarios contained in this app.

- **Neuropsychological evaluation tasks.** Included in this package, there are six apps that implement computerized versions of widely used neuropsychological tests: Dual N-Back (Jaeggi et al., 2010), Stroop task (MacLeod, 1991), Digit Span test (Ockelford, 2002), Corsi Block-Tapping test (Kessels et al., 2000) and Go/No-go task(Verbruggen and Logan, 2008). These apps alllow to record biosignals while the users are using doing the tests, allowing to study the physiological response to this work. Figure 5.13(g), 5.13h, and 5.13(i) show screenshots of three of these test.

Additionally, it is worthy to mention that MEDUSA© Platform is also designed to enable the creation of custom experiments. To make this easier, the platform provides Qt- and Unity-based templates to simplify the design and development of new apps, with a generic workflow that can be applied in most experiments. The platform's official documentation includes a comprehensive guide on how to design and develop apps for MEDUSA©, including real examples. These apps can then be shared through the official app market at https://www.medusabci.com/market.

Chapter 6

Discussion

In this dissertation, we have examined various challenges associated with current BCIs and proposed solutions to address these issues. On one hand, we focused on the development and optimization of signal processing techniques for ERP-based spellers. The goal was to improve the performance of these systems by refining the methods used to process and interpret the EEG data they rely on. Firstly, we explored the effects of non-target stimuli in these systems to propose a novel method, called OSRD, for the control state detection task. This method was tested with 15 subjects, showing a suitable asynchronous management of the speller. Afterwards, we developed EEG-Inception for P300 classification to improve the performance of the system in the command decoding task. This CNN model was trained and tested with a database of 73 subjects, including 42 severely disabled, outperforming the previous approaches. Finally, we implemented a novel method that used EEE-Inception to provide an asynchronous control of our system. This last method was tested in 22 subjects, showing outstanding results. Noteworthy, this was the first asynchronous ERP-based speller that used a fully deep-learning approach for the command decoding and control state detection tasks. On the other hand, we developed a novel software ecosystem to accelerate BCI and neuroscience research called MEDUSA©. With this platform, we aimed to simplify the development and distribution of BCI experiments and signal processing methods. In this chapter, the main outcomes of this doctoral dissertation are discussed. Additionally, the main limitations of the conducted research are presented in the last section.

6.1 SSVEP elicited by non-target stimuli from ERP-based spellers

As explained in section 1.4.1, providing a suitable asynchronous control of BCIs for communication and control in general (e.g., SSVEP-based spellers, c-VEP-based spellers), and ERP-based spellers in particular, is essential. Without it, these systems have no practical applications outside the laboratory. The first step to achieve such asynchronous control is the identification and characterization of EEG patterns that can be used for this purpose. In this dissertation, the two experiments presented in section 5.1 aimed to analyze the SSVEP triggered by non-target stimuli of an ERP-based speller using the RCP.

In the first experiment, we tested different values of SOA, as detailed in Table 5.1, to study how the variation of the stimulation rate affected the SSVEP. As showed in Figure 5.2, all the stimulation frequencies provoke a SSVEP in the user's EEG. However, it is interesting to see how this waveform has higher power for lower stimulation rates, being maximum at 2.67 Hz. On the other hand, the amplitude of the SSVEP decreased for 6.66 Hz and 8.00 Hz. Additionally, the second and third order harmonics are only clearly noticeable for 2.10 Hz, 2.67 Hz and 3.54 Hz. We hypothesize that, as non-target stimuli are received with the peripheral field of vision, higher stimulation rates are more difficult to detect, thus triggering weaker responses. This tendency is in accordance with the work of Zhang et al. (2019), who demonstrated that the SSVEPs elicited by peripheral flickering stimuli at 15 Hz have significantly lower amplitude than those received in the central region of the receptive field. This analysis facilitated our choice of SOA for the Asynchrony database. In this regard, the experiments had to be performed to maximize the power of the SSVEP to increase the accuracy of the control state detection task, without increasing too much the selection time of the system. Table 6.1 explains this trade-off, comparing the SSVEP power and the selection time using one stimulation sequence for each SOA. Despite the fact that the maximum SSVEP power is reached at $f_{\text{st}} = 2.67$ Hz (SOA = 375 ms), this frequency would increase too much the command selection speed. Comparatively, the average power of the SSVEP for a stimulation rate of $f_{\text{st}} = 5.71$ Hz (SOA = 175 ms) is comparable, but the speed of selection is doubled. Moreover, this value is similar to other related studies (He et al., 2017; Martínez-Cagigal et al., 2019a; Pinegger et al., 2015). Accordingly, we chose $f_{\text{st}} = 5.71$ Hz in this dissertation to perform the control state detection experiments and acquire the Asynchrony database.

Table 6.1: Comparative of the SSVEP power and speed of selection

SOA	Stimulation frequency	SSVEP power	Selection time
475 ms	2.10 Hz	0.007 $\mu V^2/Hz$	5.55 s
375 ms	2.67 Hz	0.010 $\mu V^2/Hz$	4.55 s
275 ms	3.64 Hz	0.008 $\mu V^2/Hz$	3.33 s
175 ms	5.71 Hz	0.009 $\mu V^2/Hz$	2.00 s
150 ms	6.66 Hz	0.004 $\mu V^2/Hz$	1.78 s
125 ms	8.00 Hz	0.005 $\mu V^2/Hz$	1.49 s

SSVEP power: averaged power of the SSVEP across all participants in the first analysis of the characterization experiment, SOA: stimulus onset asynchrony, Selection time: time required to select a command with one sequence of stimulation.

The study of the spatial distribution of SSVEPs revealed interesting patterns. At lower frequencies (2.10 Hz, 2.67 Hz, 3.64 Hz, 5.71 Hz), the power of the SSVEPs was higher in the electrodes located on the midline of the brain (i.e., Fz, Cz, CPz, Pz). Nevertheless, as the stimulation rate increased, the power of the SSVEPs moved towards electrodes situated closer to the occipital lobe, where the visual cortex is located (i.e., POz, PO7, PO8, Oz). This seems to contradict the results of previous studies, which generally find that SSVEPs are triggered by stimulation rates above 6 Hz and primarily found in the visual cortex (Bin et al., 2009; Chuan Jia et al., 2011; Friman et al., 2007; Lin et al., 2006). This variation in results may be attributed to the fundamental differences in the stimulation methodologies. Most studies use a conventional flickering cell where the participant has to focus to elicit an SSVEP. Thus, this paradigm does not involve any cognitive effort (Wolpaw and Wolpaw, 2012). In contrast, in the RCP, the participant has to distinguish between target and non-target stimuli to select a command, which engages high-level cognitive functions. At low stimulation rates, the individual may have the ability to differentiate the stimuli consciously, which may engage frontal and parietal regions of the brain, linked with complex cognitive processes and decision making (Standring, 2015). As the frequency of stimulation increases, it becomes challenging for the participant to differentiate these stimuli consciously, leaving the primary visual cortex as the most active area. Although this was the first study that examined the frequency range and spatial distribution of the SSVEP elicited by the RCP, previous research demonstrated that even minimal differences in stimulation paradigms can change the frequency range, amplitude, and brain sources of this control signal, making our hypothesis plausible (Bachiller et al., 2015; Mazaheri and Picton, 2005; Odom et al., 2010; Polich, 2007).

The findings of the second experiment, as illustrated in Figure 5.3, indicate that the SSVEP in the overt mode is evoked by the non-target stimuli of the RCP, which are perceived in the peripheral field of vision. On the other hand, the covert matrix

only triggers transient ERPs with the P300 potential. Furthermore, the averaged EEG signal during control state is the outcome of the linear superposition of these two components, whereas the non-control mode does not show any synchronized activity. This suggests that the SSVEP and ERPs evoked by the RCP have distinct underlying mechanisms and are independent.

These results shed light on the causes and characteristics of the SSVEPs elicited by the RCP, laying the groundwork for using this control signal to our advantage. Using the outcomes of these analyses, we developed two novel methods to detect the user's control state in ERP-based spellers based on the RCP. The results of the corresponding studies are discussed in the following sections.

6.2 Feature-engineering approach for control state detection: the OSRD method

Once the SSVEPs elicited by RCP were analyzed, we were able to design a successful method to detect them: the OSRD method. This method uses spectral and correlation features to detect the SSVEP in each trial. If this control signal is detected, the system assumes that the user is controlling the speller, continuing to the command decoding stage. On the contrary, if the SSVEP is not detected, the system starts a new trial without selecting any command. Thus, the OSRD method acts as an automatic switch of the system. Additionally, we proposed a novel approach to create synthetic non-control observations that reduces the calibration time, avoiding the need to register non-control trials to train the classifier. This method was tested with 15 subjects in both offline and online experiments, as detailed in section 5.2. The following sections discuss the main results of these experiments.

6.2.1 Evaluation experiment

Regarding the offline experiments, we firstly validated the synthetic non-control observations approach. As can be seen in Table 5.2, the accuracy values were similar when the model was trained with synthetic and real non-control observations. In fact, the statistical analysis did not find significant differences, regardless the number of sequences (p-value > 0.05). Thus, synthetic non-control observations are comparable to real non-control observations. The advantage of using this approach is that it reduces the duration of the calibration sessions by half, since the acquisition of non-control trials is no longer needed. Furthermore, both

approaches achieved high accuracy, with 13 out of 15 participants reaching at least 90%. Additionally, we can see that the higher the number of sequences, the higher the precision of the OSRD method. The reason is that a longer period of stimulation provokes a more discernible, stable SSVEP.

Table 5.3 shows the results of the online session. The results obtained in the offline experiment were replicated in these online sessions, achieving an average accuracy of (95.5%) for the control state detection task and demonstrating the ability of the OSRD method to work in real time. The results for PPV, NPV, TPR and TNR indicate that the method is more reliable in detecting the non-control state. The most plausible reason is the decrease in the power of the SSVEP due to momentary losses of concentration in control trials, which leads the system to detect non-control state (i.e., false negative). On the contrary, detecting a SSVEP during the non-control state (i.e., false positive), is highly unlikely given the design of the OSRD method. In this regard, a maximum performance for both classes is always desired, but the consequences of false positives and false negatives are often different. False negatives require the user to repeat the trial, which can hinder the system's agility, but it is not a critical issue. On the other hand, false positives can be especially frustrating, particularly in applications like wheelchair control, where unwanted actions could be even dangerous. Therefore, it is important to maximize the TPR. The OSRD method stands out in this task, achieving 96.7%. Additionally, Table 5.3 also includes the accuracy and ITR of the overall system, including the control state detection and command decoding stages. As shown, participants were able to control the asynchronous ERP-based speller without compromising the overall system performance, achieving an average accuracy of 92.5%.

6.2.2 Comparison with previous approaches

Table 6.2 shows a comparative between previous approaches of asynchronous ERP-based spellers and our system. This comparison should be made with caution due to differences in stimulation paradigms, and experimental settings. However, we will highlight the most important aspects. The first asynchronous ERP-based speller was proposed by Zhang et al. (2008), who designed a method that assessed the statistical differences between the output of a SVM fed with target and non-target epochs. This study achieved an ITR of 15 bits/min and a false positive rate (FPR) of 0.71 events/min with 4 CS, which suggests that the asynchrony management could be improved, especially if we take into account the implications

Table 6.2: Comparative of previous asynchronous ERP-based spellers.

Study	Asynch. method	Subj.	Ext. calib.	Acc_{csd} (%)	FPR (epm)	Acc_{ovr} (%)	ITR (bpm)
Zhang et al. (2008)	Probability analysis	4 CS	Yes	–	0.71	–	15.0
Aloise et al. (2011)	LDA threshold	11 CS	Yes	–	0.26	–	11.2
Martínez-Cagigal et al. (2017)	LDA threshold	5 CS 16 MD	Yes	–	–	95.75 84.14	–
Tang et al. (2019)	LDA threshold	4 CS	Yes	–	–	90.30	–
Aydin et al. (2018)	LDA threshold	10 CS	Yes	–	–	93.27	43.1
Martínez-Cagigal et al. (2019a)	LDA threshold	5 CS 16 MD	Yes	–	–	92.30 80.60	–
Pinegger et al. (2015)	SAM HAM	21 CS	Yes	79.5 95.5	– –	– –	– –
Santamaría-Vázquez et al. (2019a)	**OSRD**	**15 CS**	**No**	**95.5**	**0.14**	**92.5**	**12.4**

SAM: spectral analysis method; HAM: hybrid analysis method; CS: control subjects; MD: motor disabled subjects; Acc_{csd}: accuracy of the control state detection stage; FPR: false positive rate of the control state detection stage; epm: events per min, Acc_{cmd}: accuracy of the command decoding stage; ITR: information transfer rate of the overall system; bpm: bits per min; LDA: linear discriminant analysis; OSRD: oddball steady response detection.

of a false positive. In comparison, the OSRD method achieved a FPR of 0.14 events/min in the online session, showing the reliability of our approach.

From this study, different works proposed methods based on thresholds defined over the output scores of the LDA classifier used in the command decoding pipeline. The principle of operation is simple: a classifier trained to detect ERPs should score non-control epochs with lower probability than control epochs. Therefore, the system detects the control state if the probability of the ERP classifier is higher than a certain threshold that must be customized for each user. With this approach, Aloise et al. (2011) improved the accuracy of the asynchronous detection with respect to the work of Zhang et al. (2008), reaching a FPR = 0.26 events/min. However, our approach still improves this value. In Martínez-Cagigal et al. (2017, 2019a), we used the same asynchronous framework, obtaining overall system accuracies of 95.75% and 92.30% with CS, respectively. Among prior studies, Aydin et al. (2018) achieved the highest overall performance with an ITR = 43.15 bits/min. However, it is worth noting that this study used a different stimulation method, known as the region-based paradigm. This paradigm involves placing commands in distinct regions and selections are performed in two levels to increase speed, which makes it challenging to compare their results with other studies (Aydin et al., 2018).

So far, all the examined studies relied on a threshold, calculated with different

parameters derived from the output of the ERP classification stage (i.e., probability analysis, area under the ROC curve, etc), in order to discriminate the control state. However, in our own experience in Martínez-Cagigal et al. (2017, 2019a), threshold-based methods have several disadvantages. On one hand, the SNR of the ERPs is low. The scores of target and non-target epochs are usually comparable, making it necessary to repeat the stimulation for several sequences for an accurate command decoding. As a consequence, thresholds for detecting the user's control state are delicately balanced, making them vulnerable to small changes in the amplitude or latency of the ERPs. This can result in sudden significant decreases in the accuracy of these methods. In fact, they present high inter-session variability and must be recalibrated before each session with the BCI system, a time consuming procedure that affects the usability of the system and is frustrating for users. Moreover, this wrapper approach depends on the classifier of the command decoding stage. Therefore, the method must be modified if a different classifier is used. Nevertheless, it is important to note that while these disadvantages have a significant impact on the practical usability of the system, they are difficult to evaluate under controlled experimental conditions. This circumstance probably leads to overestimate the performance of threshold-based approaches in previous studies.

In contrast, control state detection methods independent of the command decoding task, such as OSRD, do not present the aforementioned disadvantages. Pinegger et al. (2015) proposed the first method for detecting the user's control state that was independent of the ERP classification stage. This study also relied on the SSVEPs triggered by the RCP as a means of identifying the user's control state. They proposed the spectral analysis method (SAM), which followed a simple approach: if the amplitude of the trial's FFT at the stimulation rate was higher than a certain threshold, the system detected control state, otherwise assuming non-control state. Despite the novelty of this method, its performance ($\text{TPR} = 88.3\%, \text{TNR} = 73.7\%, \text{ACC} = 79.5\%$, 15 sequences) suggested that this asynchrony method could be improved. Of all the approaches included in this section, only SAM can be directly compared with our method, as it uses the same phenomenon to identify the control state. This comparison shows that the OSRD method outperforms SAM in all metrics ($\text{TPR} = 94.5\%, \text{TNR} = 96.7\%, \text{ACC} = 95.5\%$). To improve their results, Pinegger et al. (2015) combined the FFT-based method with the ERP classifier scores in a hybrid approach, achieving the same average overall accuracy than the OSRD method (i.e., 95.5%), but using more stimulation sequences (15 *vs.* 10.7). Moreover, it still relies on the ERP classifi-

cation stage, having the same limitations than threshold-based approaches.

Finally, another crucial point is that, unlike OSRD, all the prior methods required non-control trials to be calibrated, which doubles the training time. Moreover, due to their reliance on the ERP classifier, they must be frequently recalibrated because of the high inter-session variability of the ERPs (He et al., 2017). This would ultimately consume a significant amount of time from the end users, reducing the feasibility of these methods in a practical context. The OSRD method solves this problem by generating synthetic non-control observations from control trials to enhance the usability of the system.

6.3 Deep-learning approach for command decoding: EEG-Inception

With the OSRD method, we had a fully asynchronous system based on feature engineering. The next step in this investigation was to improve the performance of our ERP-based speller using deep learning. For this purpose, we designed a novel CNN architecture for ERP detection, called EEG-Inception. In the following subsections, we discuss the design of EEG-Inception, and the results of the evaluation experiment with a population of 73 subjects (i.e., 42 healthy, 31 severely disabled) described in section 5.3.2.

6.3.1 Architecture design

The advancement of deep learning in recent years can be attributed to the progress made in natural language processing and computer vision (Lecun et al., 2015). As a consequence, the principles and structures that contributed to this achievement are specific to those areas and do not apply to EEG processing, hindering the development of deep-learning methods in this field (Schirrmeister et al., 2017). For the design of EEG-Inception, traditional and novel techniques from image classification and EEG processing were combined to achieve a new CNN that outperformed preceding models for ERP detection. The key features of this model will be discussed below.

A key feature of EEG-Inception is the integration of Inception modules tailored for EEG processing. The Inception architecture was originally introduced by Szegedy et al. (2015) for image processing, resulting in significant advancements in that field. However, before our study there were only a few studies that have used this architecture for EEG processing, and none for ERP detection (Lee et al.,

2020; Qiao and Bi, 2019; Yue and Wang, 2019). EEG-Inception includes two Inception modules with three branches that process the signal at different time scales (i.e. 500ms, 250ms, and 125ms) in accordance to the size of the convolutional filters (see Table 4.1) and EEG sample rate. Our experiments have shown that the inclusion of these modules leads to significant improvements when compared to single-scale methods. However, the model's performance remained consistent for different kernel sizes, and the optimal values may vary depending on the specific application. In this study, the scales were selected to maximize performance for ERP detection.

The incorporation of additional strategies into our architecture also resulted in significant performance improvements. Depthwise convolutions allow the extraction of independent spatial filters for each temporal pattern, dropout regularization reduces the over-fitting effect, and batch normalization and average pooling are useful to increase the stability of the model during training, avoiding problems such as vanishing gradients (Lawhern et al., 2018). In addition, the inclusion of the output block, which synthesizes the information extracted by previous layers in few, high-level features, maximizes the gain of the fine-tuning process.

Finally, the optimization of the hyperparameters is an important contribution that could help to develop new methods in the future. As seen in Figure 5.6(a), the selection of the activation function is a crucial factor. We evaluated Sigmoid, ReLU, and ELU functions, which are commonly used in deep learning. In this regard, ReLU is a popular choice in computer vision, but in our experiments, ELU showed superior performance. This is consistent with the findings of Schirrmeister et al. (2017) for SMR classification. To design deep-learning models for ERP detection, it is also essential to minimize over-fitting. In our tests, dropout regularization was found to be the most effective technique for reducing this effect, reaching the optimal value at 0.25. Finally, the learning rate had a significant impact in the model's performance too. As illustrated in Figure 5.6(b), the optimal value was 0.001.

6.3.2 Evaluation experiment

The evaluation experiment outlined in section 5.3.2 was designed to not only validate the effectiveness of EEG-Inception as part of the command decoding pipeline of our ERP-based speller, but also to provide a direct comparison with several well-established prior approaches. It is important to highlight that the comparison between these models would not have been feasible by analyzing the results

reported in the original studies due to varying training and testing strategies, as well as differences in tested populations. In addition to the results of the experiment, these aspects are also discussed in the following paragraphs.

Table 5.4 displays the command decoding accuracy for each model, number of fine-tuning trials, and sequences. Thus, it provides direct and fair comparison. As shown, EEG-Inception consistently achieved significantly better performance (p-value < 0.01) than EEGNet, DeepConvNet, CNN-BLSTM, xDAWN+RG, and rLDA. EEG-Inception also demonstrated higher robustness to inter-subject variability, as shown by the reduced standard deviation values, which is an important factor in severely disabled populations with different pathologies.

Regarding the ITR, EEG-Inception reached the highest value, 25.64 bits/min, as illustrated in Figure 5.7. This ITR is comparable to the range reported in related studies with healthy subjects, indicating that, with the appropriate signal processing algorithms, individuals with severe disabilities can achieve similar performance (Liu et al., 2018).

It is not surprising that a higher number of fine-tuning trials improves performance, given the high inter-subject variability of the P300 potential. More fine-tuning trials allow the models to learn subject-specific features, thus enhancing command decoding accuracy. However, this also increases the calibration time, reducing the usability if the system. Therefore, a balance between performance and calibration time should be found. In this regard, deep-learning models had a clear advantage over machine learning approaches, achieving suitable accuracy with fewer fine-tuning trials. Another factor to consider is the selection speed, which is influenced by the number of sequences. A greater number of sequences results in higher accuracy, regardless of the model, but also slows down the system. Therefore, it is essential to strike a balance between precision and speed for practical applications. In this regard, EEG-Inception offers higher accuracy with shorter selection time.

The proposed strategy to train EEG-Inception is an important contribution of the study, allowing for some valuable conclusions that are worth mentioning. Most related works, likely due to limitations in the available data, assessed the performance of new BCI classification approaches using either cross-subject evaluation (i.e. training and testing with data from different subjects) or intra-subject evaluation (i.e. training and testing with data from the same subject) (Lawhern et al., 2018; Liu et al., 2018; Schirrmeister et al., 2017). In contrast, we used a hybrid method that combined transfer learning and fine-tuning. For $N = 0$, the models were not fine-tuned, simulating a calibration-less approach for cross-

subject evaluation. For $N > 0$, the model was fine-tuned to the test subject with few calibration trials. This approach is crucial to study the true potential of deep learning for assistive BCIs, where reducing the calibration time is an important issue. In this regard, EEG-Inception is specifically engineered to take full advantage of this training approach, offering benefits for the practical implementation of ERP-based spellers in real-world applications.

Another crucial aspect is that we validated our approach with data from 31 severely disabled subjects, which are the target users of assistive ERP-based spellers. This allowed us to implement specific methods to improve the robustness of EEG-Inception against inter-subject variability, which explains the outstanding results in this particular, but important characteristic. In contrast, none of the previous studies that proposed deep-learning approaches for ERP-based spellers evaluated their models with this population (Borra et al., 2019; Cecotti and Gräser, 2011; Lawhern et al., 2018; Liu et al., 2018; Manor and Geva, 2015; Santamaría-Vázquez et al., 2019b; Schirrmeister et al., 2017). In this regard, individuals with severe disabilities pose a significant challenge due to their heterogeneity and unique characteristics. Generally, their performance is significantly lower compared with healthy subjects, making a direct comparison with previous studies difficult. However, it is essential to validate new algorithms with end users to obtain an accurate assessment of the model's performance in a real-world setting. Thus, we consider this point as one of our study's strengths.

6.4 Deep-learning approach for control state detection: EEG-Inception

After demonstrating the advantages of EEG-Inception for the command decoding task, the next goal was to improve the performance of the asynchronous management in our ERP-based speller using deep learning as well. Whereas the OSRD method reached high accuracy, we suspected that the SSVEPs elicited by non-target stimuli of the RCP were not the only EEG patterns that could be used to detect the user's control state. For instance, Martínez-Cagigal et al. (2019b) demonstrated that the attention mechanisms involved during the usage of an ERP-based speller increase the complexity of the EEG, measured with the multiscale sample entropy. Therefore, deep-learning models able to discriminate this kind of patterns should increase the performance feature-engineering methods for the detection of the user's control state, given that they automatically learn high-

level features from the input data. In this regard, EEG-Inception was a perfect candidate due to its multiscale approach, which could be particularly optimal for detecting different patterns of activity. To test this hypothesis, we designed the first study that tested a deep-learning model, EEG-Inception, for this task. The main results of this experiment, explained in section 5.4, are discussed bellow, providing a comparison with previous approaches.

6.4.1 Evaluation experiment

EEG-Inception demonstrated high performance in the control state detection task, as illustrated in Table 5.5. As can be seen, there is a trade off between performance and selection time. More stimulation sequences led to a higher confidence in the selection and reduced the impact of outliers, thereby improving the model's accuracy. As a downside, more sequences increase the selection time, reducing the speed of the system. In this regard, EEG-Inception was able to achieve accuracies above 91% with just one sequence of stimulation and $N \geq 20$, which is an outstanding result. The number of fine-tuning trials also had a positive impact on the performance at the expense of increasing the calibration time. Nonetheless, our approach can still achieve satisfactory results even with a limited number of training observations, showing the effectivity of our cross-subject transfer learning initialization. In fact, when simulating a plug-and-play system with no calibration for the test subject ($N = 0$), the model still achieved accuracies close to 90% for 10 or more stimulation sequences. In addition to its ability to perform well with no calibration for the test subject, EEG-Inception also has the ability to take advantage from the fine-tuning process ($N > 0$). This results in a significant increase in performance after a brief calibration. For example, for $N = 30$, the control state detection accuracy was 91.91% with just one sequence of stimulation and 96.95% with 15 sequences.

The confusion matrices, as shown in Figure 5.8, also offer some interesting observations. The rate of false negatives tends to be higher than the rate of false positives, particularly for lower values of N. This discrepancy could be attributed to lapses in concentration, a theory that is supported by the EEG analysis in Figure 5.9. This figure shows that the ERP had a lower amplitude, with the SSVEP also exhibiting less power.

The evaluation of the entire system, with the control state detection and command decoding stages, proved the feasibility of our approach. As seen in Table 5.6, our system demonstrated outstanding performance. A possible practical con-

figuration could be $N = 30$ and five stimulation sequences, which has an expected overall test accuracy of 91.3%. Additionally, the Figure 5.10 shows that the system achieved a maximum ITR of 35.54 bpm for $N = 30$ and one stimulation sequence, with one subject achieving 88.60 bpm. In regard to the calibration-free approach ($N = 0$), the system reached an accuracy of 85.38%, with an ITR of 13.57 bpm.

Finally, it is worth mentioning that the training/testing strategy plays an important role in leveraging the capabilities of the model. For instance, it allows to have subject specific models with very few training trials by exploiting cross-subject transfer learning. Ideally, in the initialization phase, the model learns common features across subjects to detect the target patterns in each case. Then, in the fine-tuning phase, the model particularizes these features to the specific characteristics of each subject, resulting in an improved performance, as it was demonstrated in our previous study (Santamaría-Vázquez et al., 2020b). At the same time, this method allows to take advantage from all the available data, improving its scalability for real use, where data from new subjects could be incorporated to the models to increase their performance.

6.4.2 Comparison with previous approaches

EEG-Inception achieved promising results in the control state detection task. Moreover, this was accomplished using an extensive database (the largest among related studies) that assures the robustness of our results. However, making direct comparisons with previous studies is often challenging due to significant variations in stimulation paradigms, experimental design, and sample characteristics. This section highlights the most important aspects. When results per subject were available, the statistical significance of the comparisons between different approaches was determined using the Mann-Whitney U-test, adjusting the FDR with Benjamini-Hochberg method (see section 4.5.2).

The performance of the proposed pipeline for control state detection was significantly higher than previous approaches that relied on feature engineering, as showed in Table 6.3. To simplify the comparison, here we only consider studies that proposed asynchrony methods independent of the command decoding stage. Note that wrapper threshold-based approaches were previously discussed in section 6.2.2. Pinegger et al. (2015) proposed the SAM method, reaching an average accuracy of 79.5% with 15 sequences of stimulation using spectral features, whereas EEG-Inception achieved 97.28% (p-value < 0.01), as shown in Table 5.5. Similarly, the proposed approach also outperformed the OSRD method, which was

Table 6.3: Comparative of previous asynchronous systems with independent control state detection and command decoding pipelines.

Study	Asynch. method	Subj.	Acc_{csd} (%)	Acc_{ovr} (%)	ITR (bpm)
Pinegger et al. (2015)	SAM	21 CS	79.5	–	–
Martínez-Cagigal et al. (2019b)	SampEn	10 CS	94.52	–	–
Santamaría-Vázquez et al. (2019a)	**OSRD**	**15 CS**	**95.5**	**92.5**	**12.4**
Santamaría-Vázquez et al. (2022)	**EEG-Inception**	**22 CS**	**97.28**	**96.68**	**35.54**

SAM: spectral analysis method; HAM: hybrid analysis method, SampEn: sample entropy; CS: control subjects; Acc_{csd}: accuracy for the control state detection task; ITR: information transfer rate of the overall system; bpm: bits per min; OSRD: oddball steady response detection.

based on spectral and correlation features (p-value < 0.01) (Santamaría-Vázquez et al., 2019a). The comparison with the results presented in Martínez-Cagigal et al. (2019b), which used sample entropy features, was also significant (p-value < 0.01). Notably, these differences are more pronounced when using few stimulation sequences. For instance, when using five sequences, the accuracy was increased by 10.41% and 19.60% with respect to the latter two studies.

Regarding the overall results, there are some noteworthy points to consider. Related studies achieved lower overall accuracy and ITR, which demonstrates the superiority of EEG-Inception. For instance, the maximum average ITR was 35.54 bpm (see Figure 5.10). In comparison, the highest values achieved by Zhang et al. (2008), Aloise et al. (2011) and Santamaría-Vázquez et al. (2019a) were 15.0 bpm, 11.2 bpm (p-value < 0.01) and 12.3 bpm (p-value < 0.01), respectively. On the other hand, Tang et al. (2019) obtained a maximum accuracy of 90.30% compared to 96.47% in our analysis.

In terms of calibration time, this study was the first to investigate an asynchronous ERP-based BCI without any user's calibration (i.e., $N = 0$), achieving an accuracy of 85.38% and ITR of 13.57 bpm. These results exceeded by a large margin an overall performance of 70%, which is generally considered the minimum acceptable accuracy to control a BCI (Kübler et al., 2004). In this regard, we think that minimizing, or even eliminating, the system's calibration is essential for increasing the usability of asynchronous ERP-based spellers for practical assistive applications. Thus, this is a key contribution of our work, paving the way for further research in this area.

In summary, the results presented in Santamaría-Vázquez et al. (2022) demonstrated that EEG-Inception provides great advantages for the control state detection task in ERP-based spellers. The proposed approach reached higher performance than the previous methods. It also addressed some of the main limitations of

these systems for practical assistive applications, such as the calibration time and instability of wrapper approaches based on thresholds. In this regard, it is worth noting that none of the related studies explored the possibility of a calibration-free approach.

With this study, we completed the final design of our asynchronous ERP-based speller. The system is fully based on deep learning, using EEG-Inception for the control state detection and command decoding tasks. As shown by the results presented Santamaría-Vázquez et al. (2022, 2020b), this design represents a significant step forward in the state of the art of practical asynchronous ERP-based spellers, providing higher performance than previous approaches.

6.5 Design of MEDUSA© and comparison with previous BCI platforms

MEDUSA© is a novel BCI platform that was developed to address the limitations of existing BCI software. It was created as part of this research project to help us to achieve our goals and fill the gap of available options. The goal was to accelerate BCI and cognitive neuroscience research by facilitating the implementation and open-source distribution of new systems and methods, as the ones that have been presented in this dissertation. Furthermore, MEDUSA© also includes additional BCI paradigms and EEG processing methods that were developed in parallel projects as part of the research line.

The results of this work were described in section 5.5. In this section, we discuss the most important features MEDUSA©, providing at the same time a detailed comparative with other general-purpose BCI software, such as OpenVibe or BCI2000. This comparison only includes those platforms that provide the three main stages of a BCI system: signal recording, signal processing, and feedback presentation. Thus, we leave out toolboxes for bio-signal analysis, such as MNE or EEGLAB. A summarized comparison of the various BCI platforms currently available can be found in Table 6.4.

The first aspect to discuss is the programming framework in which each platform has been developed, as this characteristic has an important impact in other features. The majority of platforms were developed in C++, including BCI2000, OpenVibe, xBCI and BF++, while Pyff was the only one implemented Python (Bianchi et al., 2003; Kothe and Makeig, 2013; Renard et al., 2010; Schalk et al., 2004; Susila et al., 2010; Venthur et al., 2010). In this regard, C++ is a general-

purpose and efficient language, but it has a steep learning curve and the development time is increased due to its complex syntax. This makes difficult to maintain the software and keep up with the latest advances in the field. On the other hand, Python is a script-like, flexible language that reduces the development time thanks to its simplicity. Additionally, it has a wide range community packages (e.g., Scipy, Numpy, Tensorflow) that simplify the development of signal processing methods. For these reasons, this language has gained popularity in the last years. The main drawback of Python is its lower efficiency compared to compiled languages like C++, but this can be addressed with careful implementations and concurrent processing for most use cases. So far, Pyff was the only BCI software that was completely written in Python, but its development has been interrupted for more than 7 years. Additionally, BCpy2000 is a Python package designed to facilitate the development of BCIs using BCI2000 as a foundation (Luczak and Marcus, 2020). However, it is important to note that BCpy2000 has important limitations in terms of functionality and carries forward other disadvantages inherited from BCI2000. This makes MEDUSA© the leading alternative implemented in Python.

Another important consideration is the maintenance of the software. As shown in Table 6.4, most of the BCI platforms are no longer being updated. Out of the software tools considered in this analysis, only OpenVibe is currently being actively developed to stay current with the latest BCI advancements, whereas BCI2000 only receives critical updates. The remaining platforms have not been updated for several years. It is important to note that creating these platforms requires extensive knowledge in software design, DevOps practices, concurrent programming, signal processing, machine learning, or graphical engines. Due to this complexity, updating and maintaining BCI software can be especially difficult, particularly given the rapid pace of progress in the field. As a result, many projects are abandoned shortly after their release. In MEDUSA©, we apply several methods to address this issue: (1) modular design that allows upgrading the different components easily without requiring major changes to the rest of the software; (2) flexible implementation in Python, a high-level programming language well-suited for rapid development of new functionalities; (3) scalable architecture prepared to accommodate additional signal processing methods or apps as needed; and (4) tools specifically designed to promote community contributions. The development of MEDUSA© has been guided by these principles to simplify the maintenance process, which aligns with our long-term vision of the project.

MEDUSA© provides a comprehensive suite of signal processing methods and

BCI experiments. For instance, it includes deep learning models and connectivity metrics for M/EEG analysis, which can be used in offline and online experiments. Furthermore, the diverse range of apps included in MEDUSA© Platform allows for the investigation of crucial aspects within the BCI and cognitive neuroscience fields. These apps facilitate the exploration of stimulus and feedback characteristics in various BCI paradigms, the comparison of classification methods in online and offline experiments, or the characterization of bio-signals, such as EEG or ECG, under different cognitive states. It is worth noting that our c-VEP speller implementation is currently the only publicly available option for investigators interested in this BCI paradigm (Martínez-Cagigal et al., 2021). In terms of functionality, OpenVibe stands out as the only platform that offers a comparable range of capabilities, whereas the other projects are significantly less advanced in this regard (see Table 6.4). Finally, it should be noted that numerous research studies conducted in recent years have tested the feasibility of these functionalities, thus providing scientific validation for our solution (Marcos-Martínez et al., 2021; Martínez-Cagigal et al., 2019b; Pérez-Velasco et al., 2022; Santamaría-Vázquez et al., 2019a, 2022, 2020a,b).

Another aspect worth mentioning is that MEDUSA© provides tools to develop and share custom apps with less effort. The tools include templates to facilitate app development with Qt and Unity, detailed tutorials, and an official app market to share apps with the community through our website. These features create the adequate environment to encourage an active participation from a multidisciplinary community. By simplifying the implementation of BCI and cognitive neuroscience experiments and increasing their reproducibility, MEDUSA© aims to attract more developers and professionals to these fields. In fact, this community's involvement is critical to the success of the project.

Table 6.4: Comparative between different platforms for BCI experiments.

Platform	Language	Last update	Analysis toolbox	Deep Learning	Connectivity analysis	LSL	Open app develop.	App market	Novel BCI paradigms	User friendly
xBCI	C++	Dec 2008	✓							
Pyff	Python	Feb 2016								✓
BF++	C++	May 2016	✓							
BCI2000	C++	Aug 2020	✓							
OpenVibe	C++	Dec 2022	✓		✓	✓	✓			✓
MEDUSA©	Python	Dec 2022	✓	✓	✓	✓	✓	✓	✓	✓

Analysis toolbox: integrated signal processing funtions suitable for offline and online analysis; Open app development: specific tools and guides to develop custom experiments; App market: community space dedicated to share apps or experiments; Novel BCI paradigms: refers to most recent BCI paradigms, such as c-VEPs or connectivity-based neurofeedback.

6.6 Limitations

Despite the positive results that were achieved in all the studies that compose this compendium, our work is not without limitations. The first limitation is related to the databases that were used to test the proposed models in this dissertation. For instance, the OSRD method and EEG-Inception were evaluated using the Asynchrony database, which was specifically designed for asynchrony studies with ERP-based spellers. However, due to the difficulty of recruiting severely disabled patients, we were only able to include healthy subjects in this database. Therefore, the results achieved in Santamaría-Vázquez et al. (2019a, 2022) may be overestimating the accuracy of both methods in a real setting with end users. The importance of this aspect is illustrated if we compare the results achieved by EEG-Inception in Santamaría-Vázquez et al. (2020b) and Santamaría-Vázquez et al. (2022). In the latter study, which did not include disabled subjects, the overall accuracy and ITR was much higher than on the first study, where the model was evaluated on 31 severely disabled subjects, even when the system was fully asynchronous and used the same command decoding pipeline. The cause is usually related to physical and cognitive problems associated with disabilities (e.g., attention deficit, involuntary movements). This demonstrates the importance of testing these methods with target users to obtain a better estimation of the model's performance in a real setting. In this regard, it should be noted that this is a common limitation in BCI studies. For this reason, the fact that our command decoding approach based on EEG-Inception was tested on target users in Santamaría-Vázquez et al. (2020b) is a great achievement. Another limiting factor related to the databases is their size, which especially affects deep learning models. Despite the fact that the sample size used in Santamaría-Vázquez et al. (2022, 2020b) was the largest among related studies, our training strategy for EEG-Inception is designed to benefit from as much data from as many subjects as possible. Therefore, a larger sample size would increase the accuracy of the models for the control state detection and command decoding tasks. In fact, more training data could help to reduce the effect overfitting, which would allow to increase the complexity of the model. Additionally, having more test subjects would contribute to more robust and generalizable results.

Regarding EEG-Inception, this model has only been tested for P300 detection, control state detection and MI decoding, achieving very satisfactory results (Pérez-Velasco et al., 2022; Santamaría-Vázquez et al., 2022, 2020b). However, EEG-Inception introduces several architectural advantages that could help to in-

crease the classification performance in other tasks as well. We think that this model has great potential in other contexts, such as SSVEP and c-VEP decoding, sleep stage scoring, or disease detection, after appropriate fine-tuning and hyperparameter optimization processes. Nevertheless, additional experiments are required to corroborate these hypotheses. Moreover, did not tested methods such as attention modules or transformers, which could help to further improve the model. Additionally, we did not apply any method to explain the features learned by EEG-Inception in the different tasks. In this regard, explainable deep-learning models could help to gain insight into brain processes through EEG and optimize the architectures, being a field of research with great potential. Therefore, future endeavors are needed to address this issue.

MEDUSA© also present limitations. As in any software under development, there may be unexpected errors in the code that need to be fixed. In this regard, we have set up communication channels to report issues and suggestions, such as the forum and the GitHub repositories. Another limitation is that certain components of the software are only compatible with Windows. Compatibility with Linux and iOS operating systems will be addressed in the future. Additionally, MEDUSA© currently lacks certain functionalities that could expand the potential applications of the software (e.g., real-time spectrogram, SSVEP speller, auditory paradigms). Nevertheless, we are firmly committed to the continuous enhancement of our solution to establish it as one of the leading BCI platforms. As a result, forthcoming versions of MEDUSA© will integrate additional functionalities, leveraging the platform's design advantages. To encourage the community's engagement and participation in the project, we will also complete the website with more features, tutorials, and documentation.

Chapter 7

Conclusions

The common thread of this doctoral dissertation has been the development of new signal processing methods to improve the control state detection and command decoding stages of ERP-based spellers, making these systems more practical for real applications outside the laboratory. In this regard, we proposed one feature-engineering method, called OSRD, to improve the asynchronous management of this BCI systems. Then, we increased the performance of our approach with the development of EEG-Inception. Additionally, while doing this work, we realized the limitations of the available BCI platforms. Therefore, we created MEDUSA© to improve the efficiency of implementing and distributing BCI systems, and increase the impact of our work.

In this chapter, the main contributions of this collection of publications are highlighted in section 7.1. Then, section 7.2 presents the joint conclusions of these studies. Finally, future research related to this work is outlined in section 7.3.

7.1 Contributions

The main contributions of this dissertation are:

1) The OSRD method: a novel feature-engineering approach that provides an asynchronous control of ERP-based spellers by detecting the SSVEP elicited by the RCP using spectral and correlation features. This is the first method that does not require to increase the calibration time of the system thanks to the creation of synthetic non-control observations from control trials. Moreover, the offline and online evaluation of the OSRD method with 15 subjects

showed its advantages over previous approaches (Santamaría-Vázquez et al., 2019a).

2) EEG-Inception: a novel deep-learning model that can be used for different EEG classification tasks. To the best of our knowledge, this is the first CNN for EEG processing that analyzes this signal from a multiscale perspective through the combination of Inception modules with other state-of-the-art structures in a lightweight architecture. These features, together with the proposed training strategy, which applies cross-subject transfer learning and fine-tuning, improved the performance of the model in comparison to existing approaches (Santamaría-Vázquez et al., 2020b).

3) The validation of EEG-Inception for the command decoding pipeline of our ERP-based speller with 73 subjects (31 of whom had motor disabilities). This evaluation showed that this model outperforms five previous approaches: rLDA, xDAWN+RG, CNN-BLSTM, DeepConvNet and EEGNet. Additionally, to the best of our knowledge, this was the first study that tested deep learning models for ERP-based spellers with data from severely disabled subjects, the target users of this technology (Santamaría-Vázquez et al., 2020b).

4) The validation of EEG-Inception for the control state detection pipeline of our ERP-based speller with 22 healthy subjects. This evaluation demonstrated, for the first time, the ability of a deep learning architecture to provide an asynchronous control of these systems, outperforming previous methods in this task. Additionally, the experiments showed the feasibility of the first asynchronous ERP-based speller fully based on deep learning, achieving high overall accuracy and ITR with few calibration trials and stimulation sequences. This solution represents a significant milestone in the field, paving the way for more practical applications of ERP-based spellers (Santamaría-Vázquez et al., 2022).

5) MEDUSA©: a novel software ecosystem optimized for BCI and cognitive neuroscience research. This innovative software ecosystem offers advanced signal acquisition functions, state-of-the-art signal processing methods, and a complete suite of ready-to-use BCI and neuroscience experiments. Furthermore, to our best knowledge, MEDUSA© provides the most comprehensive set of tools for developing and sharing new BCI applications with the community. These groundbreaking features make MEDUSA© one of the most

promising open-source platforms in the BCI field (Santamaría-Vázquez et al., 2023).

7.2 Main conclusions

The analysis and discussion of the results achieved in this work allows for the formulation of the main conclusions of this doctoral dissertation:

1) ERP-based spellers must incorporate a control state detection stage to prevent accidental command selections. This is especially important in assistive applications to allow users to be independent of supervision, thereby promoting personal autonomy.

2) The characterization of EEG patterns that relate to control or non-control states in a BCI provides valuable insights, which can be then used to implement asynchronous systems. Particularly, the RCP elicits a SSVEP at the stimulation frequency provoked by non-target stimuli presented at a fixed rate.

3) The OSRD method provides a robust control state detection in ERP-based spellers. This method is characterized by its independence from the command decoding stage, as well as its potential to achieve high performance. Moreover, it does not require to extend the calibration time of the system by utilizing synthetic non-control observations.

4) Deep-learning models outperform classical feature-engineering approaches for the command decoding and control state detection stages of ERP-based spellers. The development of network architectures and training strategies specifically tailored for EEG processing is key to improve the performance of these models.

5) EEG-Inception has been validated as a highly effective classifier for the control state detection and command decoding stages in ERP-based spellers. This CNN is characterized by the integration of Inception modules to process the EEG in different scales, its effective regularization techniques, and its efficiency in exploiting subject-specific fine-tuning with few calibration trials.

6) The performance of motor-impaired individuals is significantly lower than that of healthy subjects due to problems in attentional control, involuntary

movements, or cognitive impairment, among others. Additionally, the design of the system must take into account the special needs of this group. Therefore, assistive BCIs must be tested with their target population to ensure their practicality in a real-world setting.

7) There is a need for specialized software tools to facilitate the implementation of neurotechnology experiments, especially in the fields of BCI and cognitive neuroscience. These tools could speed up experimentation, reduce project costs, and enable the participation of researchers without the required technical knowledge to develop such software.

8) MEDUSA© implements a wide variety of tools to accelerate BCI and cognitive neuroscience research. However, in order to fully establish its position as a leading software in these domains, the involvement of a broader community of developers and content creators is essential.

7.3 Future research lines

There are several areas for future research that arise from this investigation, some of which may expand upon or build upon the findings of this dissertation, and others that may address other topics of the BCI field.

The first research line is to test our system with target users of assistive ERP-based spellers: motor-impaired subjects. It is important to validate the proposed algorithms in this population to obtain a more reliable estimation of their performance with target users. In order to pursue this research line, it would be needed to expand our *asynchrony database* with severely disabled subjects with different pathologies. Moreover, a long-term longitudinal validation in a practical setting could give insight into how socio-technological interactions and disease progression would affect the efficacy of the BCI system.

Another line of investigation could be to validate the performance of EEG-Inception on other EEG classification tasks and public datasets. This would help to determine the generalization ability of this CNN and its potential applications in a wider range of contexts, allowing for a more comprehensive understanding of the model's capabilities.

EEG-Inception could be further improved by incorporating recent developments in the deep-learning field, such as attention modules or transformers. The training strategy could be optimized too by using self-supervised strategies to leverage the large amount of unlabeled EEG data that is currently available. Moreover,

adaptive approaches for deep-learning models could lead to more efficient and effective BCI systems as well.

Another potential direction for future research is the use of explainable artificial intelligence (XAI) techniques to gain insight into brain processes through EEG data by providing interpretability and transparency in the decision-making process of a deep-learning model. These techniques can help researchers and clinicians to optimize the model architectures for BCI, interpret the results, and identify potential biases or limitations in the model's performance.

Finally, MEDUSA© is a software under active development that will continue to improve in the future. The main lines of action include: fixing bugs, improving the coverage of the platform by supporting Linux and iOS operating systems, and adding a wider variety of signal processing methods and BCI applications. Moreover, the development of more documentation, tutorials, and features to foster the involvement of the community in the project will be addressed too.

Appendix A

Compendium of publications

1. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Diego Marcos-Martínez, Víctor Rodríguez-González, Sergio Pérez-Velasco, Selene Moreno-Calderón and Roberto Hornero, "MEDUSA©: A novel Python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research", *Computer Methods and Programs in Biomedicine*, March, 2023, DOI: 10.1016/j.cmpb.2023.107357.

2. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Sergio Pérez-Velasco, Diego Marcos-Martínez, Roberto Hornero, "Robust Asynchronous Control of ERP-Based Brain-Computer Interfaces using Deep Learning", *Computer Methods and Programs in Biomedicine*, vol. 215, March, 2022, DOI: 10.1016/j.cmpb.2022.106623.

3. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Fernando Vaquerizo-Villar, Roberto Hornero, "EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-based Brain-Computer Interfaces", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28 (12), pp. 2773 - 2782, December, 2020, DOI: 10.1109/TNSRE.2020.304810.

4. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Javier Gomez-Pilar, Roberto Hornero, "Asynchronous Control of ERP-based BCI Spellers Using Steady-State Visual Evoked Potentials Elicited by Peripheral Stimuli", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27 (9), pp. 1883–1892, August, 2019, DOI: 10.1109/TNSRE.2019.2934645.

Appendix B

Scientific achievements

B.1 Publications

B.1.1 Papers indexed in the JCR

1. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Diego Marcos-Martínez, Víctor Rodríguez-González, Sergio Pérez-Velasco, Selene Moreno-Calderón and Roberto Hornero, "MEDUSA©: A novel Python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research", *Computer Methods and Programs in Biomedicine*, March, 2023, DOI: 10.1016/j.cmpb.2023.107357.

2. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Sergio Pérez-Velasco, Diego Marcos-Martínez, Roberto Hornero, "Robust Asynchronous Control of ERP-Based Brain-Computer Interfaces using Deep Learning", *Computer Methods and Programs in Biomedicine*, vol. 215, March, 2022, DOI: 10.1016/j.cmpb.2022.106623.

3. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Fernando Vaquerizo-Villar, Roberto Hornero, "EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-based Brain-Computer Interfaces", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28 (12), pp. 2773 - 2782, December, 2020, DOI: 10.1109/TNSRE.2020.304810.

4. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Javier Gomez-Pilar, Roberto Hornero, "Asynchronous Control of ERP-based BCI Spellers

Using Steady-State Visual Evoked Potentials Elicited by Peripheral Stimuli", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27 (9), pp. 1883–1892, August, 2019, DOI: 10.1109/TNSRE.2019.2934645.

5. Sergio Pérez-Velasco, **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Diego Marcos-Martínez, Roberto Hornero, "EEGSym: Overcoming Inter/subject Variability in Motor Imagery Based BCIs with Deep Learning", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1766-1775, Junio, 2022, DOI: 10.1109/TNSRE.2022.3186442.

6. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Roberto Hornero, "Brain–Computer Interface Channel Selection Optimization using Meta-heuristics and Evolutionary Algorithms", *Applied Soft Computing*, vol. 115, pp. 108176, January, 2022, DOI: 10.1016/j.asoc.2021.108176.

7. Víctor Martínez-Cagigal, Jordy Thielen, **Eduardo Santamaría-Vázquez**, Sergio Pérez-Velasco, Peter Desain, Roberto Hornero, "Brain–computer interfaces based on code-modulated visual evoked potentials (c-VEP): a literature review", *Journal of Neural Engineering*, vol. 18 (6), pp. 061002, November, 2022, DOI: 10.1088/1741-2552/ac38cf.

8. Diego Marcos-Martínez, Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Sergio Pérez-Velasco, Roberto Hornero, "Neurofeedback Training Based on Motor Imagery Strategies Increases EEG Complexity in Elderly Population", *Entropy*, vol. 23 (12), pp. 1574, November, 2021, DOI: 10.3390/e23121574.

9. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, Verónica Barroso-García, **Eduardo Santamaría-Vázquez**, Félix del Campo, David Gozal, Roberto Hornero, "A convolutional neural network architecture to enhance oximetry ability to diagnose pediatric obstructive sleep apnea", *IEEE Journal of Biomedical and Health Informatics*, vol. 25 (8), pp. 2906-2916, August, 2021, DOI: 10.1109/jbhi.2020.3048901.

10. Saúl J. Ruiz-Gómez, Roberto Hornero, Jesús Poza, **Eduardo Santamaría-Vázquez**, Víctor Rodríguez-González, Aarón Maturana-Candelas, Carlos Gómez, "A new method to build multiplex networks using Canonical Correlation Analysis for the characterization of the Alzheimer's disease contin-

uum", *Journal of Neural Engineering*, vol. 18 (2), pp. 026002, February, 2021, DOI: 10.1088/1741-2552/abd82c.

11. Marco Simões, Davide Borra, **Eduardo Santamaría-Vázquez**, GBT-UPM, Mayra Bittencourt-Villalpando, Dominik Krzemiński, Aleksandar Miladinovic, Neural-Engineering-Group, Thomas Schmid, Haifeng Zhao, Carlos Amaral, Bruno Direito, Jorge Henriques, Paulo Carvalho, Miguel Castelo-Branco, "BCIAUT-P300: A Multi-Session and Multi-Subject Benchmark Dataset on Autism for P300-Based Brain-Computer-Interfaces", *Frontiers in Neuroscience*, vol. 14, pp. 978, September, 2020, DOI: 10.3389/fnins.2020.568104.

12. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Javier Gomez-Pilar, Roberto Hornero, "Towards an Accessible Use of Smartphone-Based Social Networks through Brain-Computer Interfaces", *Expert Systems With Applications*, vol. 120, pp. 155–166, April, 2019, DOI: 10.1016/J.ESWA.2018.11.026.

13. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Roberto Hornero, "Asynchronous Control of P300-based Brain–Computer Interfaces using Sample Entropy", *Entropy*, vol. 21 (3), pp. 230, February, 2019, DOI: 10.3390/E21030230.

B.1.2 Book chapters

1. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Javier Gomez-Pilar, Roberto Hornero, "A brain–computer interface web browser for multiple sclerosis patients", in *Neurological Disorders and Imaging Physics, Volume 2: Engineering and Clinical Perspectives of Multiple Sclerosis*, ISBN-10: 0750317604, pp. 12:1–12:31, Bristol, United Kingdom: Institute of Physics Publishing, Editors: Ayman El-Baz and Jasjit S. Suri, 2019, DOI: 10.1088/978-0-7503-1762-7CH12.

B.1.3 International conferences

1. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Víctor Rodríguez-González, Sergio Pérez-Velasco, Diego Marcos-Martínez, Roberto Hornero, "MEDUSA: A Novel Platform for Modern Non-invasive Brain-computer Interfaces", *IUPESM World Congress on Medical Physics and*

Biomedical Engineering 2022, Singapur (Singapore), June 12 - June 17, 2022.

2. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Daniel Rodríguez, Jaime Finat, Roberto Hornero, "Preventing Cognitive Decline in Elderly Population through Neurofeedback Training: A Pilot Study", *5th International Conference on Neurorehabilitation (ICNR2020)*, Virtual conference (Spain), October 13 - October 16, 2020.

3. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Javier Gomez-Pilar, Roberto Hornero, "Deep learning architecture based on the combination of convolutional and recurrent layers for ERP-based brain-computer interfaces", *XV Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON 2019)*, ISBN: 978-3-030-31635-8, pp. 1844–1852, Coimbra (Portugal), September 26 - September 28, 2019, DOI: 10.1007/978-3-030-31635-8_224.

4. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Sergio Pérez-Velasco, Diego Marcos-Martínez, Selene Moreno-Calderón, Roberto Hornero, "Non-binary m-sequences for reliable, high-speed Brain–Computer Interfaces based on c-VEP: a pilot study", *IUPESM World Congress on Medical Physics and Biomedical Engineering 2022*, Singapur (Singapore), June 12 - June 17, 2022.

5. Sergio Pérez-Velasco, Gonzalo C. Gutiérrez-Tobal, Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Roberto Hornero, Assessment of Residual Deep Neural Networks and AdaBoost to predict adherence to digital-based active and healthy aging interventions, *IUPESM World Congress on Medical Physics and Biomedical Engineering 2022*, Singapur (Singapore), June 12 - June 17, 2022.

6. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Roberto Hornero, "A Portable P300-based Brain–Computer Interface as an Alternative Communication Device", *5th International Conference on Neurorehabilitation (ICNR2020)*, Virtual conference (Spain), October 13 - October 16, 2020.

7. Víctor Rodríguez-González, Jesús Poza, Pablo Núñez, Carlos Gómez, María García, Yoshihito Shigihara, Hideyuki Hoshi, **Eduardo Santamaría-Vázquez**, Roberto Hornero, "Towards Automatic Artifact Rejection in

Resting-State MEG Recordings: Evaluating the Performance of the SOUND Algorithm", 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, ISBN: 978-1-5386-1311-5, pp. 4807-4810, Berlin (Germany), July 23 - July 27, 2019.

8. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Roberto Hornero, "Controlling a Smartphone with Brain-Computer Interfaces: A Preliminary Study", *X Conference on Articulated Motion and Deformable Objects (AMDO 2018)*, ISBN: 978-3-319-94543-9, pp. 34–43, Palma de Mallorca (Spain), July 12 - July 13, 2018, DOI: 10.1007/978-3-319-94544-6_4.

9. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Roberto Hornero, "A Novel Hybrid Swarm Algorithm for P300-Based BCI Channel Selection", *World Congress on Medical Physics & Biomedical Engineering (IUPESM 2018)*, ISBN: 978-981-10-9022-6, pp. 41–45, Prague (Czech Republic), June 3 - June 8, 2018, DOI: 10.1007/978-981-10-9023-3_8.

B.1.4 National conferences

1. **Eduardo Santamaría-Vázquez**, Sergio Pérez-Velasco, Víctor Martínez-Cagigal, Roberto Hornero, "EEG-InceptionGen: Una Red Convolucional de Propósito General para la Clasificación de señales EEG", *XXXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2021)*, ISBN: 978-84-09-36054-3, pp. 163-166, Madrid (Spain), November 25 - November 26, 2021.

2. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Javier Gomez-Pilar, Roberto Hornero, "Control asíncrono de sistemas BCI basados en ERP mediante la detección de potenciales evocados visuales de estado estable provocados por los estímulos periféricos del paradigma oddball", *11º Simposio CEA de Bioingeniería (CEA 2019)*, pp. 1–11, Valencia (Spain), July 18 - July 19, 2019.

3. **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Roberto Hornero, "MEDUSA: Una Nueva Herramienta Para El Desarrollo De Sistemas Brain-Computer Interface Basada en Python", *10º Simposio CEA de Bioingeniería (CEA 2018)*, ISSN: 2341-4243, pp. 97–102, Madrid (Spain), July 2 - July 3, 2018.

4. Selene Moreno-Calderón, Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Sergio Pérez-Velasco, Diego Marcos-Martínez, Roberto Hornero, "Conecta 4: un Videojuego Multijugador para Sistemas Brain-Computer Interface basados en c-VEPs", *Jornadas de Robótica, Educación en Automática y Bioingeniería 2022 (JREB 2022)*, pp. 246-252, Málaga (Spain), May 18 - May 20, 2022.

5. Diego Marcos-Martínez, **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Sergio Pérez-Velasco, Selene Moreno-Calderón, Roberto Hornero, "ITACA: Un nuevo sistema de entrenamiento cognitivo mediante Neurofeedback", *Jornadas de Robótica, Educación en Automática y Bioingeniería 2022 (JREB 2022)*, Málaga (Spain), May 18 - May 20, 2022.

6. Selene Moreno-Calderón, Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Sergio Pérez-Velasco, Diego Marcos-Martínez, Roberto Hornero, "Evaluación de un videojuego multijugador basado en Brain Computer Interfaces utilizando c-VEPs", *XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2022)*, ISBN: 978-84-09-45972-8, pp. 324-327, Valladolid (Spain), November 23 - November 25, 2022.

7. Sergio Pérez-Velasco, Diego Marcos-Martínez, **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Selene Moreno-Calderón, Roberto Hornero, "Caracterización espacio-temporal de la clasificación de imaginación motora con herramientas de explainable artificial intelligence (XAI)", XL *Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2022)*, ISBN: 978-84-09-45972-8, pp. 316-319, Valladolid (Spain), November 23 - November 25, 2022.

8. Diego Marcos-Martínez, Ana Martín-Fernández, Sergio Pérez-Velasco, **Eduardo Santamaría-Vázquez**, Víctor Martínez-Cagigal, Selene Moreno-Calderón, Roberto Hornero, "Análisis de los cambios en la conectividad funcional tras un entrenamiento cognitivo mediante Neurofeedback", *XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2022)*, ISBN: 978-84-09-45972-8, pp. 27-30, Valladolid (Spain), November 23 - November 25, 2022.

9. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Sergio Pérez-Velasco, Diego Marcos-Martínez, Selene Moreno-Calderón, Roberto Hornero, "Un nuevo método de parada temprana no paramétrico para sistemas

Brain–Computer Interface basados en c-VEP", *XL Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2022)*, ISBN: 978-84-09-45972-8, pp. 196-199, Valladolid (Spain), November 23 - November 25, 2022.

10. Víctor Martínez-Cagigal, **Eduardo Santamaría-Vázquez**, Sergio Pérez-Velasco, Diego Marcos-Martínez, Roberto Hornero, "Sobre la Eficacia del Principio de Vecinos Equivalentes en Sistemas BCI basados en c-VEP", *12º Simposio CEA de Bioingeniería*, pp. 32-36, Madrid (Spain), June 3 - June 4, 2021.

11. Fernando Vaquerizo-Villar, Daniel Álvarez, Leila Kheirandish-Gozal, Gonzalo C. Gutiérrez-Tobal, Verónica Barroso-García, **Eduardo Santamaría-Vázquez**, Félix del Campo, David Gozal, Roberto Hornero, "Modelo de deep learning basado en la arquitectura Inception para el diagnóstico de la apnea del sueño infantil mediante la señal de oximetría", *XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2020)*, ISBN: 978-84-09-25491-0, pp. 340-343, Valladolid (Spain), November 25 - November 27, 2020.

12. **Víctor Martínez-Cagigal**, Eduardo Santamaría-Vázquez, Roberto Hornero, "Interfaz Cerebro–Ordenador para el Control de las Funcionalidades de un Teléfono Móvil", *XXXVI Congreso Anual de La Sociedad Espñola de Ingeniería Biomédica (CASEIB 2018)*, ISBN: 978-84-09-06253-9, pp. 61–64, Ciudad Real (Spain), November 21 - November 23, 2018.

13. **Víctor Martínez-Cagigal**, Javier Gomez-Pilar, Daniel Álvarez, Eduardo Santamaría-Vázquez, Roberto Hornero, "Sistema Brain-Computer Interface de Navegación Web Orientado a Personas con Grave Discapacidad", *XXXVIII Jornadas de Automática (JA 2017)*, ISBN: 978-84-16664-74-0, pp. 313–319, Gijón (Spain), September 6 - September 8, 2017.

B.2 International internship

Three-month research internship at the Spaulding Neuromodulation Center, Spaulding Rehabilitation Hospital, Harvard University, Boston, Massachusetts, USA.

i. **Purpose of the internship**
The objective of the research stay was to establish a framework for col-

laboration with the receiving group and delve into the application of EEG signals and BCI systems in the field of neurorehabilitation. Three specific objectives were proposed: 1) to apply EEG analysis techniques to identify objective biomarkers that allow evaluating the progress of the neurorehabilitation therapies applied at the Spaulding Neuromodulation Center with patients with attention deficit and hyperactivity disorder (ADHD), spinal cord injury, and chronic pain; 2) to study the potential of BCI systems for improving the effectiveness of these therapies, especially in terms of patient engagement and motivation; and 3) to develop new algorithms and software tools based on machine learning techniques for optimizing the performance of BCI systems in these applications.

ii. Summary of results

During the research stay, three studies were conducted to evaluate the effectiveness of various neuromodulation therapies. The first study was a retrospective analysis of the effectiveness of a therapy combining neurofeedback and peripheral electrical stimulation in a group of 60 children with ADHD. The second study was a preliminary evaluation of the effectiveness of a neuromodulation therapy in 5 patients with spinal cord injuries, using low-frequency laser stimulation and electrical stimulation on the affected area. The third study aimed to evaluate the potential of the placebo effect to treat patients suffering from intense chronic pain due to spinal cord injury or stroke. This therapy used a technique called operant conditioning, in which a placebo pill was administered along with the patient's prescribed analgesic medication, with the goal of gradually reducing the analgesic dosage without increasing the patient's pain. This work aimed to identify the neural mechanisms underlying the placebo effect in patients with chronic pain using EEG and fNIRs.

iii. Quality indicators of the institution

The research stay took place at the Spaulding Neuromodulation Center, under the supervision of the Associate Director of Research, Dr. Leon Morales-Quezada, MD, PhD. The center is part of the Spaulding Rehabilitation Institute at Harvard University, which is a leading global medical center in the field of physical medicine and rehabilitation. The Spaulding Neuromodulation Center is a clinical research unit dedicated to the application of innovative, non-invasive neurorehabilitation, and neuromodulation techniques. The center currently consists of about 30 people, over half of whom

hold an MD or PhD degree, and it is at the forefront of advancing the science of neurorehabilitation to treat conditions such as stroke, spinal cord injury, depression, and ADHD. Over the past 15 years, the center has studied the impact of non-invasive stimulation, developed practical clinical treatments for people with mental disorders, and promoted clinical research on brain electrical activity. Its approach is based on a multidisciplinary perspective: neuroscientific, neuropsychiatric, and rehabilitation. As part of Harvard Medical School, and with close collaborations with other institutions, the center has access to valuable databases with records of various biological signals, allowing for a deep understanding of various methodological aspects aligned with the topics of this dissertation. Dr. Morales-Quezada, the supervisor of the stay, is the Associate Director of Research at the Spaulding Neuromodulation Center Laboratory and the Director of the Clinical Neuromodulation Program at Spaulding Hospital. His curriculum includes 25 high-impact scientific publications in the field of neuroscience in the past 5 years. He is also an international consultant in pediatric neurological rehabilitation for the Happy Hope Association in Europe and Neocemod for Mexico and South America, and he is considered a pioneer in non-invasive brain stimulation research. The doctoral student also had the opportunity to work with the Director of the Spaulding Neuromodulation Center, Prof. Felipe Fregni, MD, PhD, and Professor at Harvard Medical School (h-index of 120), a world leader in the research of new neurorehabilitation therapies.

B.3 Awards and honors

- **Best bioengineering research article of the year 2020** by Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN) for the following article: Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Fernando Vaquerizo-Villar, Roberto Hornero, "EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-based Brain-Computer Interfaces", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28 (12), pp. 2773 - 2782, December, 2020, DOI: 10.1109/TNSRE.2020.304810.

- **BR41N.IO Brain-Computer Interface Designers' Hackathon** at BCI & Neurotechnology Spring School 2021 organized by g.Tec and

IEEE Brain for the project "Towards P300 calibration-less single-trial classification".

- **IMFAHE's Nodal Award 2020 in shark tank competition** organized by International Mentoring Foundation for the Advancement of Higher Education (IMFAHE) for the project:"Artifical Intelligence for Brain Cancer Diagnosis".

- **IFMBE Scientific Challenge Award** at the XV Mediterranean Conference On Medical And Biological Engineering And Computing (MEDICON2019), held at Coimbra, Portugal - September 26th-28th, 2019 for the following work: Eduardo Santamaría-Vázquez, Víctor Martínez-Cagigal, Javier Gomez-Pilar, Roberto Hornero, "Deep learning architecture based on the combination of convolutional and recurrent layers for ERP-based brain-computer interfaces", *XV Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON 2019)*, ISBN: 978-3-030-31635-8, pp. 1844–1852, Coimbra (Portugal), September 26 - September 28, 2019, DOI: 10.1007/978-3-030-31635-8_224.

Apéndice C

Resumen en castellano

C.1 Introducción

A lo largo de la historia, hemos buscado formas de liberarnos de las limitaciones del cuerpo e interactuar con el mundo directamente a través de la mente. Los sistemas *brain-computer interface* (BCI) son la materialización de esta ambición, ya que permiten a las personas controlar dispositivos externos directamente utilizando la actividad cerebral. Esta posibilidad comenzó a desarrollarse en el siglo XX gracias al progreso de la neurociencia. Los avances de las técnicas de neuroimagen permitieron estudiar nuestro sistema nervioso central con una precisión que era impensable hace apenas unas décadas, transformando por completo nuestra comprensión del cerebro y convirtiendo este campo en una de las áreas de investigación multidisciplinar más fructíferas de la actualidad. En este contexto, tanto los laboratorios de investigación como la industria están explotando este conocimiento para construir sistemas BCI, una tecnología con el potencial de transformar campos como la interacción hombre-máquina, la neurorrehabilitación y el entretenimiento, entre otros. Un sistema BCI capaz de interpretar pensamientos, deseos o intenciones podría cambiar drásticamente la forma en que interactuamos con nuestro entorno, desbloqueando un mundo inimaginable de oportunidades —y riesgos—. Aunque esta tecnología aún está muy lejos de ofrecer esta posibilidad, el campo ha evolucionado rápidamente en las últimas décadas para construir sistemas BCI cada vez más complejos y precisos.

Formalmente, un sistema BCI proporciona una vía de comunicación alternativa entre el usuario y el entorno mediante la decodificación de la actividad cerebral en tiempo real para sustituir, restaurar, potenciar, complementar o mejorar nuestras

interacciones con el entorno. En general, todos los sistemas BCI tienen un flujo de trabajo común con tres etapas: (1) registro de la actividad neuronal con técnicas de neuroimagen; (2) decodificación de los datos registrados mediante técnicas de procesado de señal y aprendizaje automático para detectar las intenciones del usuario; y (3) traducción de estas intenciones en comandos de aplicación y ejecución de los mismos. Este flujo de trabajo general se representa en la Figura 1.2. Como puede verse, estas etapas forman un bucle cerrado: las intenciones del usuario, codificadas en su actividad cerebral, desencadenan respuestas que tienen un impacto perceptible en el entorno que conducen a interacciones posteriores. Por tanto, este flujo de trabajo sigue las reglas del condicionamiento operante, una característica crucial para alcanzar una interacción natural.

El primer paso para implementar un sistema BCI es medir la actividad neuronal, lo que es un reto extremadamente complejo debido a las barreras físicas (e.g., cuero cabelludo, cráneo, meninges, etc.) que protegen el cerebro y a la complejidad de este órgano. Para medir la actividad cerebral se emplean técnicas de neuroimagen cuya precisión depende de tres aspectos importantes: la resolución espacial, la resolución temporal y la cobertura del registro. La resolución espacial se refiere al número mínimo de neuronas cuya actividad puede registrarse, la resolución temporal al evento más rápido que puede detectarse y la cobertura al volumen de cerebro que se está midiendo. El carácter invasivo de la técnica es otra consideración importante. Además, factores como el coste y la facilidad de uso también deben tenerse en cuenta para aplicaciones de BCI.

Las técnicas actuales de neuroimagen pueden dividirse en dos categorías: metabólicas y electromagnéticas. Las técnicas metabólicas detectan cambios en la respuesta hemodinámica del cerebro causados por el aumento local de la actividad neuronal al realizar una tarea cognitiva determinada. Entre ellas, se incluyen el Doppler transcraneal funcional (fTCD), la tomografía por emisión de positrones (PET), la espectroscopia funcional del infrarrojo cercano (fNIRS) y la resonancia magnética funcional (fMRI). Estos métodos, que son no invasivos o mínimamente invasivos (PET), proporcionan una buena resolución espacial y permiten una amplia cobertura cerebral. Sin embargo, su resolución temporal es muy limitada debido a la lentitud de la respuesta hemodinámica, lo que dificulta la detección de eventos breves y dispersos. Además, el equipamiento necesario no es portátil, es difícil de usar y tiene un elevado coste. Por otro lado, las técnicas de neuroimagen que miden los campos eléctricos y magnéticos generados por las neuronas incluyen los potenciales de campo local (LPF), la electrocorticografía (ECoG), la electroencefalografía (EEG) y la magnetoencefalografía (MEG). Estos métodos

tienen una excelente resolución temporal debido a la velocidad de propagación de las señales electromagnéticas. En cuanto al grado de invasión, los LPF y el ECoG requieren de la implantación de electrodos dentro del cráneo por lo que, aunque son métodos que ofrecen una excelente resolución espacial, son altamente invasivos y conllevan un riesgo elevado para el sujeto. En cuanto al EEG y la MEG, son técnicas no invasivas y completamente seguras que además proporcionan una cobertura completa del córtex mediante la colocación de electrodos sobre el cuero cabelludo. Como contrapartida, su resolución espacial es muy limitada ya que refleja la actividad conjunta de millones de neuronas al mismo tiempo. También se debe tener en cuenta que estas técnicas únicamente registran la actividad de las neuronas más superficiales del córtex, por lo que no dan información acerca de áreas más profundas. Por último, es importante destacar que, aunque la MEG es una técnica más precisa, también tiene un coste mayor y no es portátil.

Teniendo en cuenta todos los aspectos que deben considerarse para diseñar sistemas BCI, el EEG es la mejor técnica en la mayoría de los escenarios. Las razones son su excelente resolución temporal, cobertura total de la superficie del córtex, no invasividad, menor coste y usabilidad, siendo el punto más débil su baja resolución espacial. Otras técnicas que pueden utilizarse en determinados escenarios son el ECoG, que puede emplearse en casos extremos en sujetos con grave discapacidad, y la fNIRS/fMRI en algunas aplicaciones de neurorrehabilitación. Sin embargo, el EEG es, sin lugar a duda, la técnica de neuroimagen más extendida en el campo de los sistemas BCI y es la que se ha utilizado en esta investigación.

La señal registrada con el EEG se compone de actividad espontánea y evocada. La actividad espontánea está presente en ausencia de entradas o salidas explícitas y se atribuye tanto al procesamiento consciente como al inconsciente. La actividad evocada se desencadena por acontecimientos específicos, como sonidos repentinos o el movimiento de un brazo. La actividad evocada está ligada a los acontecimientos y deja una huella específica en el EEG que puede reproducirse si se repiten las condiciones del acontecimiento. Para estudiar la actividad evocada, se realizan múltiples repeticiones del mismo experimento y se promedian diferentes características del EEG para anular la influencia de la actividad espontánea. Esta operación da lugar a un potencial causado por el evento (event-related potential: ERP), que es la forma de onda característica asociada al evento estudiado. Cabe destacar que, debido a las limitaciones del EEG en cuanto a resolución espacial y cobertura de áreas profundas del cerebro, únicamente se puede identificar la huella asociada a eventos específicos que involucran a un gran número de neuronas de la superficie del córtex. Por esta razón, un sistema BCI no puede detectar en esta

señal ideas o conceptos de forma directa, sino que se tienen que aplicar técnicas de neuroingeniería para codificar en el EEG la intención del usuario y posteriormente aislar esta información mediante técnicas de procesado de señal y traducirla en un comando de aplicación.

Para codificar las intenciones del usuario en el EEG los sistemas BCI utilizan ciertos tipos de actividad evocada, denominadas señales de control. Durante las últimas décadas, los investigadores han estudiado diferentes paradigmas para optimizar el uso de la actividad evocada para el control de esta tecnología, utilizando diferentes técnicas para desencadenar respuestas adecuadas para cada aplicación. En esta tesis doctoral nos hemos centrado en los potenciales evocados P300. En este paradigma se presentan al usuario diferentes celdas en una pantalla, cada una asociada a un comando, que se iluminan secuencialmente en orden aleatorio. La tarea del usuario consiste en mirar fijamente al comando objetivo mientras se ignoran los demás estímulos. En esta configuración, los eventos asociados al comando deseado provocan ERP visuales con el potencial P300 en el EEG, mientras que los otros estímulos desencadenan ERP visuales sin esta componente, como se puede ver en la Figura 1.6. El componente P300 es un potencial positivo provocado por el reconocimiento de un estímulo poco frecuente (objetivo) dentro de una serie de estímulos frecuentes (no objetivo), que aparece alrededor de 300 ms después del inicio del estímulo. Para decodificar la orden, el sistema BCI detecta el P300 en el EEG mediante técnicas de procesamiento de señal. Esta forma de onda se considera una componente endógena, ya que refleja tareas cognitivas de alto nivel para diferenciar entre los dos tipos de estímulos, pero el usuario no necesita entrenamiento para provocarla. Esta señal de control se considera una de las más fiables para sistemas prácticos fuera del laboratorio, siendo probablemente la más extendida en el campo de los sistemas BCI en la actualidad. Los sistemas que utilizan esta señal de control se conocen como BCI basados en ERP o BCI basados en P300. Aunque ambos términos se utilizan a menudo indistintamente en la literatura, en este documento utilizamos el término BCIs basados en ERP.

Una vez que la información para discriminar la orden del usuario se ha codificado en el EEG utilizando los P300 u otras señales de control, el sistema debe detectar la actividad neuronal correspondiente y aislarla del resto de componentes del EEG aplicando métodos de procesamiento de señal. Aunque el algoritmo dependerá del paradigma específico de BCI, podemos definir un marco general de procesamiento de señales implementado en cuatro etapas secuenciales: pre-procesado, extracción de características, selección de características y clasificación de características. Una vez decodificadas las intenciones del usuario, estas se traducen en comandos

de aplicación. Un sistema BCI funciona como una interfaz general entre el cerebro del usuario y un dispositivo externo. Por tanto, esta tecnología puede utilizarse potencialmente para un número ilimitado de aplicaciones. Sin embargo, debido a su elevada complejidad y bajo rendimiento en la actualidad, su uso está restringido a campos específicos. En el momento actual las aplicaciones más importantes son los sistemas de asistencia a personas con grave discapacidad, en los que se centra esta tesis doctoral, y la neurorrehabilitación.

Aunque los sistemas de BCI basados en EEG han evolucionado enormemente en los últimos años, todavía existen varias limitaciones que deben solventarse antes de que esta tecnología esté preparada para un uso práctico generalizado como sistema de asistencia para personas con grave discapacidad. Entre ellas se pueden destacar la siguientes: (1) el registro del EEG, que requiere equipos que no están preparados para un uso continuado; (2) la baja fiabilidad y rendimiento por la alta variabilidad de la señal EEG; (3) la falta de validación con sujetos reales con grave discapacidad para adaptar los sistemas a sus necesidades concretas; (4) su funcionamiento síncrono, que impide al usuario realizar otras tareas cuando el sistema está activo, por lo que se necesita un supervisor; y (5) la complejidad de las herramientas software necesarias para la investigación y desarrollo de aplicaciones en este campo.

Las contribuciones de esta tesis doctoral se enfocaron en resolver algunas de las limitaciones anteriores, presentando un compendio de cuatro publicaciones indexadas en el *Journal Citation Reports* (JCR) entre los años 2019 y 2023. Las publicaciones se centran en: (1) mejorar el control asíncrono mediante detección de potenciales visuales de estado estable (stady-state visual evoked potencials: SSVEP) residuales (Santamaría-Vázquez et al., 2019a); (2) mejorar mediante técnicas de aprendizaje profundo la precisión en la detección de los ERP (Santamaría-Vázquez et al., 2020b); (3) mejorar la detección del estado de control del usuario mediante técnicas de aprendizaje profundo (Santamaría-Vázquez et al., 2022); y (4) desarrollar una nueva plataforma para acelerar la investigación en BCI y neurociencia cognitiva (Santamaría-Vázquez et al., 2023).

C.2 Hipótesis y objetivos

Esta tesis doctoral se centró en tres de las limitaciones que afectan a los deletreadores basados en ERP: el control síncrono; la fiabilidad y rendimiento; y las herramientas *software* para la investigación y desarrollo de sistemas BCI. Las hipótesis que guiaron cada uno de los estudios que componen el compendio se describen en

los párrafos siguientes.

En cuanto a la primera limitación, partimos de la siguiente hipótesis: *el patrón de estimulación utilizado en los sistemas BCI basados en ERP provoca diferentes tipos de actividad cerebral que pueden detectarse en el EEG*. Basándonos en el trabajo de Pinegger et al. (2015), también asumimos que *esta actividad cerebral puede utilizarse para proporcionar una detección fiable del estado de control del usuario sobre el sistema BCI*. Estas hipótesis guiaron el primer y tercer estudio del compendio.

A pesar de que existían varios trabajos que trataban sobre la aplicación de técnicas de aprendizaje profundo para el procesamiento de EEG y, más concretamente, para la detección de ERP, la mejora conseguida por estos métodos no era tan significativa como en otros campos como el reconocimiento de imágenes o el procesamiento del lenguaje natural. Este hecho nos llevó a plantear nuestra siguiente hipótesis: *el diseño y desarrollo de arquitecturas de aprendizaje profundo más complejas que tengan en cuenta la dinámica espaciotemporal específica del EEG podrían mejorar el rendimiento de estos modelos para aplicaciones BCI*. Esta afirmación guio nuestra investigación para aumentar el rendimiento de los sistemas BCI asíncronos basados en ERP para aplicaciones de asistencia a personas con gran discapacidad en el segundo y tercer estudio del compendio.

La tercera limitación que se abordó en este trabajo está relacionada con las herramientas *software* disponibles para la investigación y desarrollo de sistemas BCI. En este sentido, identificamos una serie de problemas en las plataformas BCI actuales que dificultaban la implementación y la distribución de los métodos propuestos. Este análisis condujo a la siguiente hipótesis: *el desarrollo de una nueva plataforma BCI que solventara las limitaciones identificadas podría acelerar y aumentar el impacto de la investigación en este campo*. Esta última hipótesis fue el fundamento del cuarto y último trabajo del compendio.

El objetivo principal de esta tesis fue diseñar, desarrollar y probar nuevas metodologías de procesamiento de señales para mejorar el rendimiento y la usabilidad de los BCI asíncronos basados en ERP en un contexto de asistencia. Para alcanzar este objetivo general, se propusieron los siguientes objetivos específicos:

I. Caracterizar la señal EEG durante los estados de control/no control en sistemas BCI basados en ERP con el objetivo de utilizar esta información posteriormente para mejorar el control asíncrono de estos sistemas.

II. Optimizar los métodos de procesado de señales en sistemas BCI basados en ERP para aumentar el rendimiento de la etapa de detección del estado de

control utilizando nuevas características basadas en la caracterización de los estados de control/no control.

III. Aumentar el rendimiento de los sistemas BCI basados en ERP para aplicaciones de asistencia a personas con grave discapacidad utilizando técnicas de aprendizaje profundo.

IV. Desarrollar una novedosa plataforma BCI para facilitar el diseño, implementación y distribución de experimentos y aplicaciones BCI que pudiera ser utilizada a lo largo de esta investigación, permitiéndonos alcanzar los objetivos anteriores.

V. Difundir los principales resultados de este trabajo en revistas indexadas JCR, congresos nacionales e internacionales.

C.3 Sujetos

En el curso de esta investigación se utilizaron tres bases de datos independientes con datos EEG de sujetos utilizando un sistema BCI basado en ERP. Concretamente, el paradigma de estimulación utilizado fue el paradigma fila-columna (row-column paradigm: RCP). El RCP es un paradigma que muestra una matriz de comandos, cuyas filas y columnas se iluminan aleatoriamente. El usuario selecciona un comando mirando fijamente la opción deseada, lo que provoca un ERP con el componente P300 cuando se percibe el estímulo objetivo. A continuación, el sistema decodifica la fila y la columna utilizando algoritmos de procesado de señal para detectar la componente P300 y ejecutar el comando correspondiente, proporcionando realimentación al usuario. Este sistema también se conoce como deletreador BCI basado en ERP, ya que la matriz de comandos típica se corresponde con un teclado, aunque puede presentarse como cualquier otra interfaz de control. A continuación, se detallan las características de estas bases de datos y los sujetos que las componen:

1 **Base de datos de asincronía.** Esta base de datos se adquirió para caracterizar las diferencias en el EEG cuando el sujeto está controlando el paradigma RCP vs cuando está realizando otra tarea. Esta base de datos incluye señales de 22 sujetos de control (CS) (15 hombres, 7 mujeres, edad media: 24,7 $\pm$ 4,3 años) que realizaron 120 selecciones con el sistema BCI (60 de control y 60 de no control). Las señales se registraron utilizando un equipo EEG de 16 canales. Hay que tener en cuenta que los primeros 15 sujetos

Cuadro C.1: Resumen de las bases de datos utilizadas.

Base de datos	CS	MD	Paradigma	SD	ISI	Canales
Asincronía	22	0	RCP	75	100	16
Navegador Web BCI	10	15	RCP	62.5	$\mathcal{U}(125, 250)$	8
Redes sociales BCI	10	16	RCP	62.5	$\mathcal{U}(125, 250)$	8

CS: sujetos control; MD: sujetos con discapacidad motora; RCP: row-column paradigm; SD: duración del estímulo en ms, ISI: intervalo entre estímulos en ms; $\mathcal{U}$ distribución aleatoria uniforme.

fueron adquiridos para Santamaría-Vázquez et al. (2019a). Posteriormente, esta muestra se amplió para Santamaría-Vázquez et al. (2022) con 7 sujetos adicionales.

2 **Base de datos de Navegador Web BCI.** Esta base de datos se adquirió para validar un sistema BCI basado en ERP que usaba el RCP para ayudar a personas con discapacidad severa a navegar por internet con un navegador web BCI adaptado Martínez-Cagigal et al. (2017). La base de datos contiene señales de 10 CS (6 hombres, 4 mujeres, edad media: 24,8 ± 2,9) y 15 sujetos con discapacidad motora (MD) (10 hombres, 5 mujeres, edad media: 42,7 ± 7,5 años) que realizaron diferentes tareas con el navegador web. Cada sujeto realizó 87,9 ± 7,3 selecciones. Hay que tener en cuenta que algunas tareas contenían momentos de uso libre, de ahí la variabilidad en el número de ensayos por sujeto. Las señales se registraron utilizando 8 canales de EEG (Martínez-Cagigal et al., 2017).

3 **Base de datos de Redes Sociales BCI.** Esta base de datos contiene datos de un estudio de viabilidad para evaluar una aplicación BCI asistida basada en el RCP que permitía utilizar varias redes sociales en un smartphone (Martínez-Cagigal et al., 2019a). La base de datos contenía datos de 10 CS (8 hombres, 2 mujeres, edad media: 26,10 ± 3,45) y 16 MD (10 hombres, 6 mujeres, edad media: 45,5 ± 9,68 años). Cada sujeto realizó 63,4 ± 8,2 selecciones utilizando un equipo EEG de 8 canales (Martínez-Cagigal et al., 2019a).

En la tabla C.1 se resumen las características de cada base de datos. Los sujetos MD de los estudios "Navegador Web BCI" y "Redes Sociales BCI" fueron reclutados por el Centro de Referencia Estatal de Discapacidad y Dependencia, ubicado en la provincia de León.

C.4 Métodos

En esta sección se describen los métodos que se han aplicado para realizar esta tesis doctoral. En primer lugar, se explican los métodos de procesado de la señal, incluyendo el pre-procesado y las etapas de la extracción, selección y clasificación de características que han aplicado sobre las señales EEG. Sólo se incluyen las técnicas que forman parte de los métodos propuestos en la tesis doctoral, no incluyéndose aquellos que se utilizaron a efectos comparativos. Tras los métodos de procesado de señal se detallan las métricas de rendimiento, la metodología estadística y las estrategias de validación que se utilizaron en este trabajo. Por último, se explican los métodos y estrategias para desarrollar nuestra plataforma BCI.

En el análisis de señales EEG, el pre-procesado es un paso importante que ayuda a limpiar y preparar los datos brutos para su posterior análisis. Las señales de EEG suelen contener ruido y artefactos, como interferencias eléctricas, actividad muscular y movimientos oculares. Las técnicas de pre-procesado ayudan a eliminar o reducir estos artefactos, mejorando la calidad y fiabilidad general de los datos de EEG. En este trabajo se aplicaron dos técnicas de pre-procesamiento para mejorar las señales de EEG: filtrado frecuencial y filtrado espacial. Con el filtrado frecuencial se eliminaron las frecuencias fuera de la banda de interés, que fueron diferentes en cada estudio. Posteriormente, se aplicó un filtro espacial de referencia de media común (CAR) para mejorar la detección de ERP y aumentar la resolución espacial de la señal. Este método permite reducir los artefactos comunes a todos los electrodos, como las interferencias de potencia.

En el campo del aprendizaje automático, las características son propiedades medibles de un fenómeno que sirven como datos de entrada al modelo de clasificación. En el ámbito del EEG, las características más básicas son los valores de amplitud de la señal en cada electrodo en función del tiempo. Posteriormente, estos pueden transformarse para obtener información más significativa para el análisis. La elección de las características es crucial en el desarrollo de modelos EEG, ya que tiene un impacto significativo en el rendimiento y la precisión del método. En esta investigación, se utilizaron características basadas en análisis temporal, espectral y de correlación. Las primeras estaban directamente relacionadas con los valores de amplitud del EEG, tras aplicar un proceso de decimado y enventanado de esta señal para extraer épocas de un segundo de duración por cada estímulo del RCP (Santamaría-Vázquez et al., 2019a, 2020b). Las características espectrales se basaban en el cálculo de un ratio de potencia entre dos bandas de frecuencia del EEG para detectar picos a la frecuencia de estimulación del RCP

(Santamaría-Vázquez et al., 2019a). Por último, las características extraídas a partir del método de análisis de correlación canónica (CCA) permitían detectar patrones de actividad estables a una determinada frecuencia (Santamaría-Vázquez et al., 2019a). Únicamente en el primer estudio del compendio se aplicó un proceso de selección de 60 características mediante el algoritmo de regresión paso-a-paso (SW) con criterios de inclusión y exclusión: $p < 0{,}10$ y $p > 0{,}15$, respectivamente (Santamaría-Vázquez et al., 2019a).

La fase de clasificación de características de un sistema BCI utiliza algoritmos de aprendizaje automático para detectar patrones en los datos de EEG y decodificar las intenciones del usuario. En esta tesis, se utilizaron dos modelos basados en aprendizaje supervisado por su mayor robustez. Como técnica de aprendizaje automático clásico, utilizamos el análisis discriminante lineal (LDA), un algoritmo de clasificación que calcula un hiperplano de separación entre las dos clases maximizando la distancia interclase y minimizando la varianza intraclase (Bishop and Nasrabadi, 2006). Por otro lado, diseñamos y desarrollamos una nueva red convolucional (CNN) llamada EEG-Inception, basada en técnicas de aprendizaje profundo. Esta CNN fue propuesta para tareas de clasificación de EEG. Se compone de múltiples capas que extraen representaciones jerárquicas de los datos de entrada. El modelo incorpora conceptos del campo de la visión por ordenador, como los módulos Inception, para realizar un análisis multiescala de la señal de entrada, teniendo en cuenta la estructura espaciotemporal del EEG para mejorar su procesamiento. El modelo también incorpora otras técnicas como *batch normalization*, funciones de activación no lineales y regularización *dropout* para aumentar su rendimiento. La arquitectura de EEG-Inception se divide en tres bloques principales: el primer módulo realiza un primer análisis temporal y espacial por separado, el segundo módulo realiza un segundo análisis que combina toda la información temporal y espacial disponible, y el último bloque de salida se encarga de la clasificación final de la señal de entrada (Santamaría-Vázquez et al., 2020b).

La validación de los métodos propuestos se realizó a través de distintas métricas, incluyendo la precisión en la detección del estado de control, la precisión en la decodificación de comandos y la tasa de transferencia de información (ITR). También se aplicaron test estadísticos para realizar comparaciones equitativas entre grupos de resultados y extraer conclusiones del análisis. En esta tesis se han utilizado el test de Wilcoxon y la prueba U de Mann-Whitney para comparaciones pareadas y no pareadas, respectivamente (Narsky and Porter, 2013). Cuando se realizaron múltiples comparaciones, se aplicó la técnica de Benjamini-Hochbert para corregir la tasa de falsos descubrimientos (FDR) (Benjamini and Hochberg,

1995). También se utilizó *leave-one-subject-out* (LOO) como técnica de validación cruzada para evaluar la capacidad de generalización de los resultados obtenidos por los métodos propuestos.

Por último, las técnicas que se usaron para desarrollar MEDUSA©, nuestra nueva plataforma de BCI para acelerar la investigación en BCI y neurociencia son: (1) modularidad, con una arquitectura formada por dos componentes independientes: MEDUSA© Kernel y MEDUSA© Platform; (2) flexibilidad, ya que la plataforma permite implementar experimentos BCI y algoritmos de procesado de señal de forma rápida y eficiente; (3) escalabilidad, gracias a su diseño específico que permite incrementar las funcionalidades fácilmente; y (4) desarrollo en Python, un lenguaje de programación de código abierto muy popular tanto en la investigación como en la industria por su facilidad de uso y la gran cantidad de herramientas y bibliotecas especializadas desarrolladas por la comunidad, como SciPy, Numpy, Scikit-learn y Tensorflow, para procesamiento de datos, aprendizaje automático y aprendizaje profundo.

C.5 Resultados y discusión

En primer lugar se analizó la señal de EEG en cinco sujetos de control durante los estados de control y no control del paradigma RCP con el objetivo de encontrar características que permitieran discriminar con precisión estos estados y así proporcionar un control asíncrono de nuestro sistema BCI (Santamaría-Vázquez et al., 2019a). Concretamente, el objetivo era investigar los mecanismos subyacentes y las características del SSVEP provocado por patrón de estimulación del RCP partiendo del estudio de Pinegger et al. (2015), que observó por primera vez este fenómeno. El estudio consistió en dos experimentos. El primero examinó cómo variaban las características del SSVEP en función de la frecuencia de estimulación, probando seis valores de frecuencia. Los resultados mostraron que con tasas de estimulación más bajas se obtenía una mayor potencia del SSVEP. El estudio también analizó la distribución espacial de estos potenciales encontrando que, a frecuencias más bajas, la potencia del SSVEP era mayor en los electrodos situados en la línea media de las regiones frontal y parietal. Sin embargo, a medida que aumentaba la frecuencia, la potencia se desplazaba hacia los electrodos más cercanos al lóbulo occipital. En cuanto al segundo experimento, este tenía como objetivo investigar el origen de la SSVEP utilizando diferentes matrices de estimulación y diferentes modos de estímulos visuales. El análisis de los datos obtenidos permitió concluir que el fenómeno estudiado está provocado por los estímulos no objetivo

del paradigma de estimulación recibidos a través del campo visual periférico. A partir de los resultados obtenidos en estos dos experimentos, se estableció la frecuencia óptima de funcionamiento del RCP (i.e., 5.71 Hz) y la distribución de los electrodos para los experimentos de detección del estado de control y la adquisición de la *base de datos de asincronía* (Santamaría-Vázquez et al., 2019a).

Basándonos en los resultados de los experimentos anteriores, propusimos un nuevo método basado en ingeniería de características denominado *oddball stedy-state response detection* (OSRD) para detectar el estado de control del usuario en el RCP (Santamaría-Vázquez et al., 2019a). El método OSRD detecta el SSVEP provocado por el patrón de estimulación de este paradigma, proporcionando una salida binaria $y \in \{0, 1\}$ que se corresponde con los estados de no control y control, actuando como interruptor automático del sistema. El método OSRD incluye filtrado FIR de paso de banda, filtrado espacial CAR y extracción de características basadas en análisis espectral y de correlación, que luego se clasificaron mediante LDA. También se propuso un enfoque novedoso para crear observaciones sintéticas de no control, lo que reduce el tiempo de calibración. El método se probó con 15 sujetos en experimentos *offline* y *online*. En el experimento *offline* se validó el enfoque basado en observaciones sintéticas. Posteriormente, el rendimiento del método se evaluó en la sesión online, alcanzando una alta precisión final del 95,5 % y una ITR de 12.4 bits/min. Los resultados también mostraron que el método es más fiable a la hora de detectar el estado de no control, lo que es importante para evitar acciones no deseadas en aplicaciones como el control de sillas de ruedas. El rendimiento global del sistema, incluidas las fases de detección del estado de control y decodificación de comandos fue también satisfactorio, alcanzando una precisión media del 92,5 %. Los resultados se compararon con enfoques anteriores en deletreadores asíncronos basados en ERP, mostrando que el método OSRD los supera en términos de precisión, velocidad de selección y tiempo de calibración.

Con el método OSRD disponíamos de un sistema totalmente asíncrono basado en extracción y clasificación de características. El siguiente paso en esta investigación fue mejorar el rendimiento de nuestro deletreador basado en ERP mediante técnicas aprendizaje profundo (Santamaría-Vázquez et al., 2020b). Para ello, propusimos una arquitectura CNN para el procesado de señales EEG, denominada EEG-Inception, con el fin de mejorar la precisión de decodificación de comandos de nuestro deletreador asíncrono basado en ERP. Las principales características de EEG-Inception eran la integración de módulos Inception, combinados con técnicas de procesado de imagen y EEG. El modelo también aplica técnicas como convoluciones *depthwise*, regularización *dropout*, *batch normalization* y *average pooling*,

con una estructura especialmente diseñada para evitar el efecto de sobreentrenamiento. Además, la optimización de los hiperparámetros y la metodología de entrenamiento de la red también fueron contribuciones destacables del estudio. El conjunto de datos utilizado para la evaluación de EEG-Inception estaba formado por las tres bases de datos presentadas en la sección C.3, contando con un total de 42 sujetos sanos y 31 sujetos con grave discapacidad. Los sujetos sanos se asignaron a los conjuntos de entrenamiento y validación, mientras que los sujetos con grave discapacidad se asignaron al conjunto de test. El modelo propuesto se validó simulando un escenario realista, donde la cantidad de datos de entrenamiento de un único sujeto es limitada, y teniendo en cuenta que los usuarios reales de los deletreadores basados en ERP son personas con grave discapacidad. El modelo se comparó con cinco métodos anteriores (i.e., EEGNet, DeepConvNet, CNN-BLSTM, xDAWN+RG, y rLDA), mostrando un rendimiento en el conjunto de test significativamente mayor en términos de precisión (84.6 %) y ITR (25.64 bits/min). Para interpretar correctamente estos resultados en comparación con los estudios anteriores, debe tenerse en cuenta que el modelo se evaluó en un grupo heterogéneo de 31 sujetos con grave discapacidad. Esta población presenta una gran variabilidad provocada por problemas físicos y cognitivos asociados a las distintas patologías (e.g., déficit de atención, movimientos involuntarios), lo que demuestra la importancia de probar estos métodos en esta población para obtener una mejor estimación del rendimiento del modelo en un entorno real. En este sentido, EEG-Inception demostró una mayor robustez a la variabilidad intersujeto que los anteriores modelos.

Tras demostrar las ventajas de EEG-Inception para la tarea de decodificación de comandos en los sistemas BCI basados en ERP, el siguiente objetivo era mejorar el rendimiento del control asíncrono en nuestro deletreador RCP utilizando también técnicas de aprendizaje profundo. Mientras que el método OSRD alcanzó una alta precisión, sospechábamos que los SSVEP provocados por la estimulación del RCP no eran los únicos patrones EEG que podían utilizarse para detectar el estado de control del usuario. Por ejemplo, en Martínez-Cagigal et al. (2019b) demostramos que los mecanismos de atención implicados en el uso del sistema aumentaban la complejidad del EEG, medida con la entropía muestral multiescala. Por tanto, los modelos de aprendizaje profundo, capaces de discriminar este tipo de patrones mediante la optimización automática de características, podrían aumentar el rendimiento de los anteriores métodos. En este sentido, EEG-Inception era un candidato perfecto debido a su arquitectura y enfoque multiescala, que podría ser especialmente óptimo para detectar distintos patrones de actividad. Por esta

razón, propusimos un nuevo método que empleaba este modelo para la tarea de detección del estado de control en deletreadores ERP (Santamaría-Vázquez et al., 2022). Este enfoque fue validado con la *base de datos de asincronía*, siguiendo una metodología similar al estudio de validación de EEG-Inception. Los resultados alcanzados fueron muy favorables, mostrando que EEG-Inception era capaz de incrementar la precisión (e.g., más del 90 % de precisión con una secuencia de estimulación), la velocidad de selección y el tiempo de calibración en comparación con las propuestas anteriores, incluido el método OSRD. En este sentido, cabe destacar que este fue el primer estudio que implementó un sistema BCI basado en ERP asíncrono completamente basado técnicas de aprendizaje profundo, por lo que los resultados alcanzados en esta tesis doctoral representan un avance significativo del estado del arte de los deletreadores RCP.

Durante el desarrollo de la investigación presentada en los estudios anteriores, nos dimos cuenta de que las herramientas *software* disponibles para la investigación de la BCI no se adaptaban totalmente a nuestras necesidades. Por esta razón, decidimos crear una nueva plataforma BCI llamada MEDUSA© implementada en Python (Santamaría-Vázquez et al., 2023). Este trabajo se realizó de manera consistente a lo largo de todo el desarrollo de la tesis doctoral. El objetivo inicial era implementar las funciones necesarias para desarrollar y probar nuestro sistema RCP. Sin embargo, a medida que avanzaba la investigación y desarrollábamos proyectos en paralelo, el objetivo se hizo más ambicioso. Se incluyeron nuevos métodos de procesado de señal, paradigmas BCI y experimentos de neurociencia cognitiva, así como funcionalidades para fomentar un investigación más transparente y colaborativa en este campo. En concreto, MEDUSA© Kernel es una librería que incluye un amplio conjunto de métodos para el procesamiento de señales, entre los que se incluyen filtros frecuenciales, filtros espaciales, métricas de activación local, métricas de conectividad y modelos para paradigmas BCI basados en ERP, paradigmas MI, paradigmas basados en c-VEP y neurofeedback, incluyendo todos los presentados en este documento como OSRD o EEG-Inception. Por otro lado, MEDUSA© Platform es un programa de escritorio con una arquitectura flexible y adaptable, que puede utilizarse para una amplia gama de experimentos de BCI y neurociencia. Su arquitectura consta de tres módulos: adquisición de señales, gráficos en tiempo real y aplicaciones. Además, las funcionalidades de estos módulos se pueden controlar a través de una intuitiva interfaz gráfica. El módulo de adquisición de señales utiliza el protocolo *lab-streaming layer* (LSL) para implementar funciones avanzadas de adquisición de datos, como la compatibilidad con cualquier dispositivo biomédico y la capacidad para gestionar múltiples entradas

de datos a la vez. El módulo de gráficos en tiempo real implementa visualizaciones en tiempo y frecuencia para representar las señales recibidas a través del módulo de adquisición. Por último, las aplicaciones son componentes que ponen en funcionamiento paradigmas BCI al tiempo que proporcionan realimentación en tiempo real y monitorizan una o varias señales a la vez. El diseño de este módulo permite la independencia entre las aplicaciones y otros módulos, posibilitando el desarrollo de nuevos componentes sin precisar cambios en el código base, lo que aumenta la escalabilidad de nuestra solución. La plataforma admite aplicaciones basadas en Qt y Unity para manejar la parte gráfica de la aplicación y actualmente ofrece seis aplicaciones listas para usar: *recorder*, deletreador RCP, deletreador c-VEP, Paradigma de imaginación motora, Neurofeedback y pruebas de evaluación neuropsicológica. Con estas funcionalidades y los métodos de desarrollo utilizados, MEDUSA© es una de las plataformas BCI más completas que existen en la actualidad, siendo la única alternativa viable en Python.

C.6 Conclusiones

El análisis y la discusión de los resultados anteriores permiten formular las principales conclusiones de esta tesis doctoral:

1) Los sistemas BCI basados en ERP deben incorporar una etapa de detección del estado de control para evitar la selección accidental de comandos. Esto es especialmente importante en las aplicaciones asistenciales para permitir a los usuarios manejar el sistema sin supervisión, fomentando así la autonomía personal.

2) La caracterización de los patrones de EEG relacionados con los estados de control o no control en un sistema BCI proporciona información valiosa que puede utilizarse para implementar sistemas asíncronos. En particular, el RCP provoca un SSVEP a la frecuencia de estimulación del sistema.

3) El método OSRD proporciona una detección robusta del estado de control en deletreadores basados en ERP. Este método se caracteriza por su independencia de la etapa de descodificación de comandos, así como por su potencial para alcanzar un alto rendimiento. Además, este método no requiere ampliar el tiempo de calibración del sistema gracias al uso de observaciones sintéticas.

4) Los modelos basados en aprendizaje profundo superan a los enfoques clásicos de ingeniería de características en las etapas de detección del estado de

control y decodificación de comandos de los sistemas BCI basados en ERP. El desarrollo de arquitecturas de red específicamente adaptadas al procesado de EEG es clave para mejorar el rendimiento de estos modelos.

5) La red EEG-Inception ha sido validada como un clasificador altamente eficaz para las etapas de detección del estado de control y decodificación de comandos en deletreadores basados en ERP. Esta CNN se caracteriza por la integración de módulos de Inception para procesar el EEG en diferentes escalas, sus eficaces técnicas de regularización y su eficiencia para aprovechar la ganancia de un entrenamiento con pocas observaciones de calibración.

6) El rendimiento de las personas con grave discapacidad es significativamente inferior al de los sujetos sanos. Esto es debido a problemas de control atencional, movimientos involuntarios o deterioro cognitivo, entre otros. Por lo tanto, los sistemas BCI asistenciales deben probarse con usuarios reales para garantizar su viabilidad en un entorno real.

7) Hacen falta herramientas *software* especializadas que faciliten la realización de experimentos de neurotecnología, especialmente en los campos de BCI y la neurociencia cognitiva. Estas herramientas podrían acelerar la experimentación, reducir los costes de los proyectos y permitir la participación de investigadores sin los conocimientos técnicos necesarios para desarrollar dichas herramientas.

8) MEDUSA© implementa una amplia variedad de herramientas para acelerar la investigación en el campo de los sistemas BCI y la neurociencia cognitiva. Sin embargo, su consolidación como plataforma líder en estos ámbitos estará ligada a la participación de una comunidad más amplia de desarrolladores y creadores de contenidos.

Bibliography

Abásolo, D., Hornero, R., Espino, P., Álvarez, D., Poza, J., 2006. Entropy analysis of the EEG background activity in Alzheimer's disease patients. Physiological Measurement 27 (3), 241–253.

Acharya, J. N., Hani, A. J., Cheek, J., Thirumala, P., Tsuchida, T. N., oct 2016. American Clinical Neurophysiology Society Guideline 2: Guidelines for Standard Electrode Position Nomenclature. Neurodiagnostic Journal 56 (4), 245–252.

Aloise, F., Schettini, F., Aricò, P., Leotta, F., Salinari, S., Mattia, D., Babiloni, F., Cincotti, F., 2011. P300-based brain-computer interface for environmental control: An asynchronous approach. Journal of Neural Engineering 8 (2).

Amiri, S., Fazel-Rezai, R., Asadpour, V., 2013. A Review of Hybrid Brain-Computer Interface Systems. Advances in Human-Computer Interaction 2013, 1–8.

Angelidis, A., Hagenaars, M., van Son, D., van der Does, W., Putman, P., 2018. Do not look away! Spontaneous frontal EEG theta/beta ratio as a marker for cognitive control over attention to mild and high threat. Biological Psychology 135 (March), 8–17.
URL https://doi.org/10.1016/j.biopsycho.2018.03.002

Aydin, E. A., Bay, O. F., Guler, I., 2018. P300-Based Asynchronous Brain Computer Interface for Environmental Control System. IEEE Journal of Biomedical and Health Informatics 22 (3), 653 – 663.

Bachiller, A., Romero, S., Molina, V., Alonso, J. F., Mañanas, M. A., Poza, J., Hornero, R., 2015. Auditory P3a and P3b neural generators in schizophrenia: An adaptive sLORETA P300 localization approach. Schizophrenia Research 169 (1-3), 318–325.

Barachant, A., Bonnet, S., Congedo, M., Jutten, C., 2012. Multiclass brain-computer interface classification by Riemannian geometry. IEEE Transactions on Biomedical Engineering 59 (4), 920–928.

Baumgartner, R. W., 2006. Handbook on neurovascular ultrasound. Vol. 21. Karger Medical and Scientific Publishers.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological) 57 (1), 289–300.

Berger, H., dec 1929. Über das Elektrenkephalogramm des Menschen. Archiv für Psychiatrie und Nervenkrankheiten 87 (1), 527–570.
URL http://link.springer.com/10.1007/BF01797193

Bianchi, L., Babiloni, F., Cincotti, F., Salinari, S., Marcian, M. G., 2003. Introducing BF++: A C++ framework for cognitive bio-feedback systems design. Methods of information in medicine 42 (01), 104–110.

Bin, G., Gao, X., Yan, Z., Hong, B., Gao, S., 2009. An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method. Journal of Neural Engineering 6 (4).

Bishop, C. M., Nasrabadi, N. M., 2006. Pattern recognition and machine learning. Springer.

Blankertz, B., Müller, K. R., Krusienski, D. J., Schalk, G., Wolpaw, J. R., Schlögl, A., Pfurtscheller, G., Millán, J. D. R., Schröder, M., Birbaumer, N., 2006. The BCI competition III: Validating alternative approaches to actual BCI problems. IEEE Transactions on Neural Systems and Rehabilitation Engineering 14 (2), 153–159.

Borra, D., Fantozzi, S., Magosso, E., 2019. Convolutional Neural Network for a P300 Brain-Computer Interface to Improve Social Attention in Autistic Spectrum Disorder. XV Mediterranean Conference on Medical and Biological Engineering and Computing–MEDICON 2019: Proceedings of MEDICON 2019, 1837–1843.

Bronzino, J. D., Peterson, D. R., 2014. Biomedical Engineering Fundamentals.

Cecotti, H., Gräser, A., 2011. Convolutional neural networks for P300 detection with application to brain-computer interfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (3), 433–445.

Chollet, F., 2015. Keras.
URL https://keras.io

Chuan Jia, Xiaorong Gao, Bo Hong, Shangkai Gao, 2011. Frequency and Phase Mixed Coding in SSVEP-Based Brain–Computer Interface. IEEE Transactions on Biomedical Engineering 58 (1), 200–206.

Ciliberti, D., Kloosterman, F., 2017. Falcon: A highly flexible open-source software for closed-loop neuroscience. Journal of Neural Engineering 14 (4), 1–16.

Congedo, M., Barachant, A., Bhatia, R., 2017. Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review. Brain-Computer Interfaces 4 (3), 155–174.

Craik, A., He, Y., Contreras-Vidal, J. L., 2019. Deep learning for electroencephalogram (EEG) classification tasks: A review. Journal of Neural Engineering 16 (3), 1–28.

Cruz, A., Pires, G., Lopes, A., Carona, C., Nunes, U. J., 2021. A Self-Paced BCI with a Collaborative Controller for Highly Reliable Wheelchair Driving: Experimental Tests with Physically Disabled Individuals. IEEE Transactions on Human-Machine Systems 51 (2), 109–119.

Delorme, A., Makeig, S., mar 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134 (1), 9–21.
URL https://linkinghub.elsevier.com/retrieve/pii/S0165027003003479

Durand, D. M., sep 2006. What is Neural Engineering? Journal of Neural Engineering 4 (4).

Enriquez-Geppert, S., Huster, R. J., Herrmann, C. S., 2017. EEG-neurofeedback as a tool to modulate cognition and behavior: A review tutorial. Frontiers in Human Neuroscience 11 (February), 1–19.

Farwell, L. A., Donchin, E., 1988. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. Electroencephalography and Clinical Neurophysiology 70 (6), 510–523.

Fornito, A., Zalesky, A., Bullmore, E. T., 2016. Fundamentals of Brain Network Analysis. Academic Press.

Friman, O., Volosyak, I., Axel, G., 2007. Multiple Channel Detection of Steady-State Visual Evoked Potentials for Brain-Computer Interfaces. IEEE Transactions On Biomedical Engineering 54 (4), 742–750.

Fu, R., Tian, Y., Bao, T., Meng, Z., Shi, P., 2019. Improvement Motor Imagery EEG Classification Based on Regularized Linear Discriminant Analysis. Journal of Medical Systems 43 (6), 1–13.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., Hämäläinen, M., 2013. MEG and EEG data analysis with MNE-Python. Frontiers in Neuroscience 7 (7 DEC).
URL http://journal.frontiersin.org/article/10.3389/fnins.2013.00267/abstract

Guyon, I., Elisseeff, A., 2003. An Introduction to Variable and Feature Selection Isabelle. Journal ofMachine Learning Research 703 (3), 1157–1182.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T. E., 2020. Array programming with NumPy. Nature 585 (7825), 357–362.

He, B., 2020. Neural Engineering.

He, S., Zhang, R., Wang, Q., Chen, Y., Yang, T., Feng, Z., Zhang, Y., Shao, M., Li, Y., 2017. A P300-Based Threshold-Free Brain Switch and Its Application in Wheelchair Control. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (6), 715–725.

Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. 32nd International Conference on Machine Learning, ICML 2015 1, 448–456.

Jacobson, S., 2008. Neuroanatomy for the neuroscientist, 1st Edition. Springer, New York.

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., Meier, B., 2010. The concurrent validity of the N-back task as a working memory measure. Memory 18 (4), 394–412.

Jessen, K. R., 2004. Glial cells. International Journal of Biochemistry and Cell Biology 36 (10), 1861–1867.

Jobson, J. D., 1991. Applied Multivariate Data Analysis. Springer Texts in Statistics. Springer New York, New York, NY.
URL http://link.springer.com/10.1007/978-1-4612-0955-3

Kessels, R. P., Van Zandvoort, M. J., Postma, A., Kappelle, L. J., De Haan, E. H., 2000. The Corsi Block-Tapping Task: Standardization and normative data. Applied Neuropsychology 7 (4), 252–258.

Kingma, D. P., Ba, J. L., 2015. Adam: A method for stochastic optimization. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.

Kothe, C., 2014. Lab Streaming Layer Protocol.
URL https://labstreaminglayer.org

Kothe, C. A., Makeig, S., 2013. BCILAB: A platform for brain-computer interface development. Journal of Neural Engineering 10 (5), 1–17.

Krusienski, D. J., Sellers, E. W., McFarland, D. J., Vaughan, T. M., Wolpaw, J. R., 2008. Toward enhanced P300 speller performance. Journal of Neuroscience Methods 167 (1), 15–21.

Krzanowski, W., 2000. Principles of multivariate analysis. OUP Oxford.

Kübler, A., Neumann, N., Wilhelm, B., Hinterberger, T., Birbaumer, N., 2004. Predictability of brain-computer communication. Journal of Psychophysiology 18 (2/3), 121–129.

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., Lance, B. J., 2018. EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. Journal of Neural Engineering 15 (056013), 1–17.

Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.

Ledoit, O., Wolf, M., 2004. A well-conditioned estimator for large-dimensional covariance matrices. Journal of Multivariate Analysis 88 (2), 365–411.

Lee, D. Y., Jeong, J. H., Shim, K. H., Kim, D. J., 2020. Classification of Upper Limb Movements Using Convolutional Neural Network with 3D Inception Block. 8th International Winter Conference on Brain-Computer Interface, BCI 2020.

Li, Y., Pan, J., Wang, F., Yu, Z., 2013. A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control. IEEE Transactions on Biomedical Engineering 60 (11), 3156–3166.

Lin, Z., Zhang, C., Wu, W., Gao, X., 2006. Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs. IEEE Transactions on Biomedical Engineering 53 (12), 2610–2614.

Liu, M., Wu, W., Gu, Z., Yu, Z., Qi, F. F., Li, Y., 2018. Deep learning based on Batch Normalization for P300 signal detection. Neurocomputing 275, 288–297.

Lopes, G., Monteiro, P., 2021. New Open-Source Tools: Using Bonsai for Behavioral Tracking and Closed-Loop Experiments. Frontiers in Behavioral Neuroscience 15 (1), 1–9.

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., Yger, F., 2018. A review of classification algorithms for EEG-based brain-computer interfaces: A 10 year update. Journal of Neural Engineering 15 (3).

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., Arnaldi, B., 2007. A review of classification algorithms for EEG-based brain-computer interfaces. Journal of Neural Engineering 4 (2).

Luck, S. J., 2014. An introduction to the event-related potential technique. MIT press.

Luczak, N., Marcus, M., 2020. BCpy2000.
URL https://github.com/neurotechcenter/BCpy2000

Luke Campagnola, 2020. PyQtGraph - Scientific Graphics and GUI Library for Python.
URL https://www.pyqtgraph.org/http://www.pyqtgraph.org/

MacLeod, C. M., 1991. Half a century of research on the Stroop effect: An integrative review. Psychological Bulletin 109 (2), 163–203.

Manor, R., Geva, A. B., 2015. Convolutional Neural Network for Multi-Category Rapid Serial Visual Presentation BCI. Frontiers in Computational Neuroscience 9 (146), 1–12.

Marcos-Martínez, D., Martínez-Cagigal, V., Santamaría-Vázquez, E., Pérez-Velasco, S., Hornero, R., 2021. Neurofeedback training based on motor imagery strategies increases EEG complexity in elderly population. Entropy 23 (12), 1–19.

Martínez-Cagigal, V., 2020. Toward Practical P300-based Brain–Computer Interfaces: Asynchrony, Channel Optimization and Assistive Applications. Ph.D. thesis, Universidad de Valladolid.
URL https://uvadoc.uva.es/handle/10324/47516

Martínez-Cagigal, V., Gomez-Pilar, J., Álvarez, D., Hornero, R., 2017. An Asynchronous P300-Based Brain-Computer Interface Web Browser for Severely Disabled People. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (8), 1332–1342.

Martínez-Cagigal, V., Santamaría-Vázquez, E., Gomez-Pilar, J., Hornero, R., 2019a. Towards an accessible use of smartphone-based social networks through brain-computer interfaces. Expert Systems with Applications 120, 155–166.

Martínez-Cagigal, V., Santamaría-Vázquez, E., Hornero, R., 2019b. Asynchronous Control of P300-Based Brain–Computer Interfaces Using Sample Entropy. Entropy 21 (3), 1–14.

Martínez-Cagigal, V., Thielen, J., Santamaría-Vázquez, E., Pérez-Velasco, S., Desain, P., Hornero, R., 2021. Brain-computer interfaces based on code-modulated visual evoked potentials (c-VEP): A literature review. Journal of Neural Engineering 18 (6), 1–21.

Mazaheri, A., Picton, T. W., 2005. EEG spectral dynamics during discrimination of auditory and visual targets 24, 81–96.

Musso, F., Brinkmeyer, J., Mobascher, A., Warbrick, T., Winterer, G., 2010. Spontaneous brain activity and EEG microstates. A novel EEG/fMRI analysis approach to explore resting-state networks. NeuroImage 52 (4), 1149–1161.

Narsky, I., Porter, F. C., 2013. Statistical analysis techniques in particle physics: Fits, density estimation and supervised learning. John Wiley & Sons.

Naseer, N., Hong, K. S., 2015. fNIRS-based brain-computer interfaces: A review.

Ockelford, A., 2002. The magical number two, plus or minus one: Some limits on our capacity for processing musical information. Musicae Scientiae 6 (2), 185–219.

Odom, J. V., Bach, M., Brigell, M., Holder, G. E., McCulloch, D. L., Tormene, A. P., Vaegan, 2010. ISCEV standard for clinical visual evoked potentials (2009 update). Documenta Ophthalmologica 120 (1), 111–119.

Panicker, R. C., Puthusserypady, S., Pryana, A. P., Sun, Y., 2010. Asynchronous P300 BCI: SSVEP-based control state detection. European Signal Processing Conference 58 (6), 934–938.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Pérez-Velasco, S., Santamaría-Vázquez, E., Martínez-Cagigal, V., Marcos-Martínez, D., Hornero, R., 2022. EEGSym: Overcoming Inter-subject Variability inMotor Imagery Based BCIs with Deep Learning. IEEE Transactions on Neural Systems and Rehabilitation Engineering 30, 1766 – 1775.

Pfurtscheller, G., McFarland, D. J., jan 2012. BCIs That Use Sensorimotor Rhythms. In: Brain–Computer InterfacesPrinciples and Practice. Oxford University Press, pp. 228–240.

Pinegger, A., Faller, J., Halder, S., Wriessnegger, S. C., Müller-Putz, G. R., 2015. Control or non-control state: that is the question! An asynchronous visual P300-based BCI approach. Journal of Neural Engineering 12 (1), 014001.

Polich, J., oct 2007. Updating P300: An integrative theory of P3a and P3b. Clinical Neurophysiology 118 (10), 2128–2148.

Proakis, J. G., 2001. Digital signal processing: principles algorithms and applications.

Qiao, W., Bi, X., 2019. Deep spatial-temporal neural network for classification of EEG-based motor imagery. ACM International Conference Proceeding Series, 265–272.

Raghavan, M., Fee, D., Barkhaus, P. E., 2019. Generation and propagation of the action potential. In: Handbook of Clinical Neurology. Vol. 160. pp. 3–22.

Renard, Y., Lotte, F., Gibert, G., Congedo, M., Maby, E., Delannoy, V., Bertrand, O., Lécuyer, A., 2010. OpenViBE : An Open-Source Software Platform to Design , Test , and Use Brain – Computer Interfaces in Real and Virtual Environments. Presence 19 (1), 35–53.

Roc, A., Pillette, L., Mladenovic, J., Benaroch, C., N'Kaoua, B., Jeunet, C., Lotte, F., 2021. A review of user training methods in brain computer interfaces based on mental tasks. Journal of Neural Engineering 18 (1), 1–35.

Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., Faubert, J., 2019. Deep learning-based electroencephalography analysis: A systematic review. Journal of Neural Engineering 16 (5).

Santamaría-Vázquez, E., Martínez-Cagigal, V., Gomez-Pilar, J., Hornero, R., 2019a. Asynchronous Control of ERP-Based BCI Spellers Using Steady-State Visual Evoked Potentials Elicited by Peripheral Stimuli. IEEE transactions on neural systems and rehabilitation engineering 27 (9), 1883–1892.

Santamaría-Vázquez, E., Martínez-Cagigal, V., Gomez-Pilar, J., Hornero, R., 2019b. Deep Learning Architecture Based on the Combination of Convolutional and Recurrent Layers for ERP-Based Brain-Computer Interfaces. XV Mediterranean Conference on Medical and Biological Engineering and Computing–MEDICON 2019: Proceedings of MEDICON 2019, 1844–1852.

Santamaría-Vázquez, E., Martínez-Cagigal, V., Hornero, R., 2018. MEDUSA: Una nueva herramienta para el desarrollo de sistemas Brain-Computer Interface basada en Python. Cognitive Area Networks 5 (1), 87–92.

Santamaría-Vázquez, E., Martínez-Cagigal, V., Marcos-Martínez, D., Rodríguez-González, V., Pérez-Velasco, S., Moreno-Calderón, S., Hornero, R., 2023. MEDUSA©: A novel Python-based software ecosystem to accelerate brain-computer interface and cognitive neuroscience research. Computer Methods and Programs in Biomedicine 230, 107357.

Santamaría-Vázquez, E., Martínez-Cagigal, V., Pérez-Velasco, S., Marcos-Martínez, D., Hornero, R., 2022. Robust Asynchronous Control of ERP-Based Brain-Computer Interfaces using Deep Learning. Computer Methods and Programs in Biomedicine 215 (1), 1–10.

Santamaría-Vázquez, E., Martínez-Cagigal, V., Rodríguez, D., Finat, J., Hornero, R., 2020a. Preventing Cognitive Decline in Elderly Population Through Neurofeedback Training: A Pilot Study. In: International Conference on NeuroRehabilitation. Springer, pp. 407–411.

Santamaría-Vázquez, E., Martinez-Cagigal, V., Vaquerizo-Villar, F., Hornero, R., 2020b. EEG-Inception: A Novel Deep Convolutional Neural Network for Assistive ERP-based Brain-Computer Interfaces. IEEE Transactions on Neural Systems and Rehabilitation Engineering 28 (12), 2773–2782.

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., Wolpaw, J. R., 2004. BCI2000: A general-purpose brain-computer interface (BCI) system. IEEE Transactions on Biomedical Engineering 51 (6), 1034–1043.

Schettini, F., Aloise, F., Aricò, P., Salinari, S., Mattia, D., Cincotti, F., 2014. Self-calibration algorithm in an asynchronous P300-based brain-computer interface. Journal of Neural Engineering 11 (3).

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., Ball, T., 2017. Deep learning with convolutional neural networks for EEG decoding and visualization. Human brain mapping 38 (11), 5391–5420.

Sellers, E. W., Arbel, Y., Donchin, E., jan 2012. BCIs That Use P300 Event-Related Potentials. In: Brain–Computer InterfacesPrinciples and Practice. Oxford University Press, pp. 215–226. URL https://academic.oup.com/book/1700/chapter/141292091

Sitaram, R., Lee, S., Birbaumer, N., jan 2012. BCIs That Use Brain Metabolic Signals. In: Brain–Computer InterfacesPrinciples and Practice. Oxford University Press, pp. 302–314. URL https://academic.oup.com/book/1700/chapter/141295063

Srinivasan, R., jan 2012. Acquiring Brain Signals from Outside the Brain. In: Brain–Computer InterfacesPrinciples and Practice. Oxford University Press, pp. 106–122. URL https://academic.oup.com/book/1700/chapter/141284904

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research 15, 1929–1958.

Standring, S., 2015. Gray's anatomy e-book: the anatomical basis of clinical practice. Elsevier Health Sciences.

Susila, I. P., Kanoh, S., Miyamoto, K. I., Yoshinobu, T., 2010. xBCI: A generic platform for development of an online BCI system. IEEJ Transactions on Electrical and Electronic Engineering 5 (4), 467–473.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1–9.

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., Leahy, R. M., 2011. Brainstorm: A User-Friendly Application for MEG/EEG Analysis. Computational Intelligence and Neuroscience 2011, 1–13. URL http://www.hindawi.com/journals/cin/2011/879716/

Tang, J., Liu, Y., Jiang, J., Yu, Y., Hu, D., Zhou, Z., 2019. Toward Brain-Actuated Mobile Platform. International Journal of Human-Computer Interaction 30 (10), 846–858.

Tantum, W. O., 2021. Handbook of EEG Interpretation. Springer Publishing Company.

Tariq, M., Trivailo, P. M., Simic, M., 2018. EEG-Based BCI Control Schemes for Lower-Limb Assistive-Robots. Frontiers in Human Neuroscience 12 (August).

Treder, M. S., Schmidt, N. M., Blankertz, B., 2011. Gaze-independent brain-computer interfaces based on covert attention and feature attention. Journal of Neural Engineering 8 (6).

Vansteensel, M. J., Jarosiewicz, B., 2020. Brain-computer interfaces for communication, 1st Edition. Vol. 168. Elsevier B.V. URL http://dx.doi.org/10.1016/B978-0-444-63934-9.00007-X

Venthur, B., Scholler, S., Williamson, J., Dähne, S., Treder, M. S., Kramarek, M. T., Müller, K.-R., Blankertz, B., 2010. Pyff – A Pythonic Framework for Feedback Applications and Stimulus Presentation in Neuroscience. Frontiers in Neuroscience 4 (1), 1–17.

Verbruggen, F., Logan, G. D., 2008. Automatic and Controlled Response Inhibition: Associative Learning in the Go/No-Go and Stop-Signal Paradigms. Journal of Experimental Psychology: General 137 (4), 649–672.

Virtanen, P., Gommers, R., Oliphant, T., Haberland, M., al. Et, 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods. Nature.Com 17 (3), 261–272.

Welch, P. D., jun 1967. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. IEEE Transactions on Audio and Electroacoustics 15 (2), 70–73.

Wolpaw, J., Wolpaw, E. W., 2012. Brain-computer interfaces: principles and practice. OUP USA.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., Vaughan, T. M., 2002. Brain Computer Interfaces for communication and control. Clinical neurophysiology 4 (113), 767–791.

Yu, Y., Zhou, Z., Liu, Y., Jiang, J., Yin, E., Zhang, N., Wang, Z., Liu, Y., Wu, X., Hu, D., 2017. Self-paced operation of a wheelchair based on a hybrid brain-computer interface combining motor imagery and P300 potential. IEEE Transactions on Neural Systems and Rehabilitation Engineering 25 (12), 2516–2526.

Yuan, K., Chen, C., Wang, X., Chu, W. C. W., Tong, R. K. Y., 2021. Bci training effects on chronic stroke correlate with functional reorganization in motor-related regions: A concurrent eeg and fmri study. Brain Sciences 11 (1), 1–16.
URL https://doi.org/10.3390/brainsci

Yue, K., Wang, D., 2019. EEG-based 3D visual fatigue evaluation using CNN. Electronics (Switzerland) 8 (11), 1–18.

Zhang, H., Guan, C., Wang, C., 2008. Asynchronous P300-based brain-computer interfaces: a computational approach with statistical models. IEEE transactions on bio-medical engineering 55 (6), 1754–63.

Zhang, N., Liu, Y., Yin, E., Deng, B., Cao, L., Jiang, J., Zhou, Z., Hu, D., 2019. Retinotopic and topographic analyses with gaze restriction for steady-state visual evoked potentials. Scientific Reports 9 (1), 1–10.

Zhang, Z., Sabuncu, M. R., 2018. Generalized cross entropy loss for training deep neural networks with noisy labels. Advances in Neural Information Processing Systems, 8778–8788.

Index

A

accuracy 59
assistive technologies 2, 8, 14, 24
asynchrony 4, 5, 25, 67, 92, 94

B

BCI platforms 6, 62, 82, 105
brain-computer interfaces 2, 9
  applications 24
  control signals 20
  limitations 25
  signal acquisition 10

C

code-modulated visual evoked potentials 21
command decoding 98
command decoding accuracy 59, 75, 100
common average reference 47
conclusions 113
contributions 111
control state detection 4, 5, 92, 101
control state detection accuracy 59, 72, 79, 94, 95, 103
cross-validation 61
  leave-one-out 61

D

databases 39, 41
deep learning 4, 5, 27, 52, 98, 101
  convolutional neural networks 28
  recurrent neural networks 28

E

electroencephalography 2, 17, 18
  acquisition 41
  evoked activity 18
  spontaneous activity 18
  volume conduction 18
event-related desynchronization 20
event-related synchronization 20

F

feature clasification
  EEG-Inception 101
feature classification 24, 52
  EEG-Inception 4, 5, 54, 98
  linear discriminant analysis 53
feature engineering 94
feature extraction 23, 47
  correlation-based features 49, 69
  frequency-based features 48, 69
  time-based features 48, 55, 69
feature selection 23, 51
  backward elimination 51
  forward selection 51
  stepwise regression 51
filtering

frequency 46, 69, 77, 82
spatial 47, 69, 77, 82

H

hypothesis 33

I

information transfer rate 60, 72, 75, 79, 95, 100, 104

L

limitations 109
linear discriminant analysis 24

M

machine learning 52
MEDUSA 6, 62, 82, 105
motor disabilities 40, 72, 101, 109

N

negative predictive value 72
neural activity 10
neural engineering 8
neurofeedback 25
neurorehabilitation 25

O

objectives 34
oddball paradigm 22, 37
OSRD 67, 94
overall accuracy 59, 72, 79, 95, 104

P

P300 evoked potentials 22, 37
positive predictive value 72
positive predictive valuey 59, 60
pre-processing 22, 45

R

row-column paradigm 37

S

sensorymotor rhythms 20
signal processing 7, 22
spatial resolution 14
statistical analysis 60
steady-state visual evoked potentials 21, 92

T

temporal resolution 14
thematic consistency 2
true negative rate 72
true positive rate 72