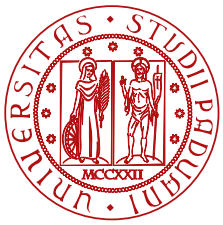


UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

DEPARTMENT OF INFORMATION ENGINEERING

MASTER'S DEGREE IN BIOENGINEERING

Development of a non invasive self-paced BCI for movement intention detection and control of a lower limb exoskeleton

Supervisor
Dr. Tortora Stefano

Candidate
Lo Faro Alessio

ACADEMIC YEAR 2024-2025

Graduation date 04/12/2025

Contents

List of Figures

Chapter 1

Introduction

1.1 A new medical challenge

The last decades have been characterized by a steady increase in the incidence of persons affected by a total, or partial, paralysis condition related to strokes, physical traumas, spinal lesions or other neurological diseases. Using Europe as the population pool, it has been found that stroke events have been the major cause of death and the second cause of acquired paralysis. A recent study compounded an incidence of 191,9 cases per 100.000 persons-year with a rising prevalence now estimated at 9.2% [1]. The projection study conducted in 2017 estimates that the number of people in the European Union which will experience a stroke event in their live, will increase around 27% over a 30 years period (2047). Reasons behind the findings are the improving survival rate of stroke events coupled by the aging trend of the population [2]. This ever increasing amount of patients will result in an unmet need for new and improved rehabilitation strategies. To these day, any recovery requires intensive physiotherapy sessions characterized by a high resource need. For this category of patients the total or partial loss of motor abilities of upper or lower limbs implies not only reduced mobility capabilities, but also lower daily life autonomy and psychological drawbacks.

1.1.1 Exoskeletons

The most immediate approach to improving the daily life of patients affected by the aforementioned conditions is the use of assistive technologies designed to compensate the lost functions. Among these, exoskeletons represent a valid option to face the deprivation of motor capabilities of both upper and lower limbs, enabling natural and coordinated movements [3]. More specifically, in the field of deambulation, lower limb exoskeletons (LLE) support the user and allow patients with severe disabilities to stand up and walk again thereby providing functional inde-

pendence that would be impossible with other assistive technologies, such as wheelchairs[4]. Beyond the purely functional advantages associated to the maintenance of a standing position, the capability of performing anatomical movements may have significant social implications. Patients may indeed experience reduced social isolation along with greater confidence in public spaces arising from a better participation in daily activity compared to wheelchair users. Taking social factors into account is crucial, since mobility limitations can contribute to social marginalization and have a negative impact on psychological well-being [5]. In this context, ongoing developments are increasingly directed toward lighter, more compact, and portable exoskeletons, aiming to maximize user comfort and achieve a level of discretion that allows them to be worn unobtrusively under clothing. [6], [7].

To face the challenges posed by the ever growing number of patients requiring assistance after paralysis-inducing events or diseases, robotic exoskeletons are also emerging as valid devices for motor rehabilitation. With low external assistance, exoskeletons allow the user to perform intensive, repetitive, and task-oriented exercises that require lower intervention by a specialist, thus facilitating the simultaneous treatment of more patients. To these topics many papers have been published, for example, in a systematic literature review of 'robot-assisted gait rehabilitation' for stroke survivors in subacute phase (≤6 months), Calafiore et al [8] analyzed 14 different randomized trials comparing the results of rehabilitation performed traditionally and with the adoption of an exoskeleton. Although the author fails to find statistical differences in the principal evaluation scales, it concludes that robot-assisted training is effective for gait recovery purposes albeit it is not superior to conventional treatments. Even more promising are other reviews which suggest that chronic stroke patients treated with an exoskeleton achieved better gait-, equilibrium- and endurance-related outcomes in the majority of the analyzed papers [9].

As supported by [10], in the long run, it is therefore easy to envision the possibility of amplifying the volume of treated patients by a single structure and reducing the cost of a single rehabilitation session without the tradeoff of a worse outcome.

1.1.2 Coupling of BCI and Exoskeletons

If interpreted purely as motion replacement devices, exoskeleton still necessitates a valid user-machine interface to control the device. Traditionally, control strategies relied on buttons, joysticks or vocal commands which in many cases may limit the naturalness, responsiveness and adaptability of the movements. Specifically, to overcome these limitations, Brain-Computer Interfaces (BCI) stand as a highly potential control layer bypassing any physical actuator and directly converting the user motor intention into a machine command, bridging the gap between intention and actuation [11].

Even more of interest is the combination of exoskeletons with BCIs, which represents a new frontier in motor rehabilitation. By controlling the exoskeleton through the patient's neural activity, this approach aims to promote self-paced, voluntary movement, thereby potentially improving motor recovery and patient engagement. Pioneer studies have already been conducted comparing the outcomes of post-stroke upper limb BCI-assisted function recovery; relevant clinical effects have been observed relatively to the control groups[12]–[14]. These studies comprise a variety of BCI paradigms with different feedback typologies (mediated for example by display cues, sensory stimulation, virtual reality visors or robotics activation) [12]. Diving deeper into the field of paired BCI and exoskeleton-mediated rehabilitation, Frolov et al. [15] performed a randomized, controlled, and multi-center trial evaluating the performances of a group of patients affected by severe upper limb paralysis. The subjects, treated in 10 session using a BCI-controlled grasp exoskeleton, were compared to the control group for which the activation of the mechanical device was independent from the brain activity of the user. Their results suggest that "adding BCI control to exoskeleton-assisted physical therapy can improve post-stroke rehabilitation outcomes" [15]. Moreover, they observed a correlation between classifier accuracy and performance improvement.
The underlying mechanism inducing a similar improvement can be identified into the causal sensory feedback provided by the activation of the external actuator as a consequence of the voluntary movement intent [16].

1.2 Thesis objectives

It should be noted that most of the existing literature has focused on upper-limb motor impairment. On the other hand, the literature on brain-driven LLE control is scarce. The present thesis aims to develop a BCI system for LLE control triggered by lower-limb motor imagery. As it will be discussed, the reproducible identification of stable BCI features, an essential prerequisite for building a fast-reacting and reliable system, remains an open challenge [17].
Within this context, the objective of the thesis is to develop and test, on healthy participants, a procedural framework that facilitates familiarization with a BCI-driven LLE. The ultimate goal is to enable control of the exoskeleton using the most natural command: a motor imagery task corresponding to a full step, performed without any actual muscular activation.
In addition to the development and testing of the physical system itself, this work also aims to analyze and interpret the experimental data collected during all sessions. Specifically, the recorded data set will be used to assess the overall feasibility of the proposed framework, as well as to explore potential indicators of user learning and adaptation over time. By examining neural and behavioral correlates, the study seeks to identify how participants progressively improve their

ability to control the system through motor imagery. Moreover, the thesis will include a comparative analysis of existing state-of-the-art deep learning architectures for EEG-based decoding. These models will be evaluated both in terms of their classification performance on the newly acquired dataset and their generalization capability across subjects and sessions. The objective is not only to benchmark model accuracy but also to determine if some architecture prove to be more effective then others for the specific task of gait motor imagery.

Chapter 2

Theory background

2.1 Brain Computer Interface (BCI)

In recent years, Brain Computer Interfaces (BCIs) have been a major topic of studies due to their high potential in the field of neurorehabilitation medicine and post-traumatic life support capabilities. This technology aims at establishing a new path of communication between a human brain and an external actuator bypassing the human peripheral nervous system and all the usual outputs it can produce, such as limb movements or speech production [18]. This capability makes BCIs uniquely suited for users who lack voluntary motor control.

A modern precursor of contemporary BCI systems is represented by the first neurofeedback loop trial of 1962 [19]. This system produces a physical feedback, such as visual auditory or tactile stimuli, proportional to preselected features of the subject's real-time measured electroencephalographic (EEG) signal. The basic principle is that, by reacting to the feedback, the subject can learn to modulate consciously its cerebral activity and consequently learn to instruct a robotic device. The main drawback of this setup is that only the subject is adapting to the machine and actively learning, and it is in this context that neurofeedback loops differentiate themselves from BCIs.

As a matter of fact, in the current view as represented in Figure 2.1, a complete BCI is described as a three-body entity comprising an active learning user, an adaptable decoding machine and an intelligent robotic actuator [18]. The described architecture enables the process of mutual learning between the subject and the system since adaptation occurs in both the subject and the decoder algorithm. The first, thanks to the feedback produced by a monitor or any robotic actuator, learns to modulate his cerebral activity to produce more stable and recognizable patterns over time. In the meantime, the second through the operator actions is subjected to constant updates across trials and sessions and consequently adjusts its parameters to discern user-imposed signals with increasingly higher accuracy.

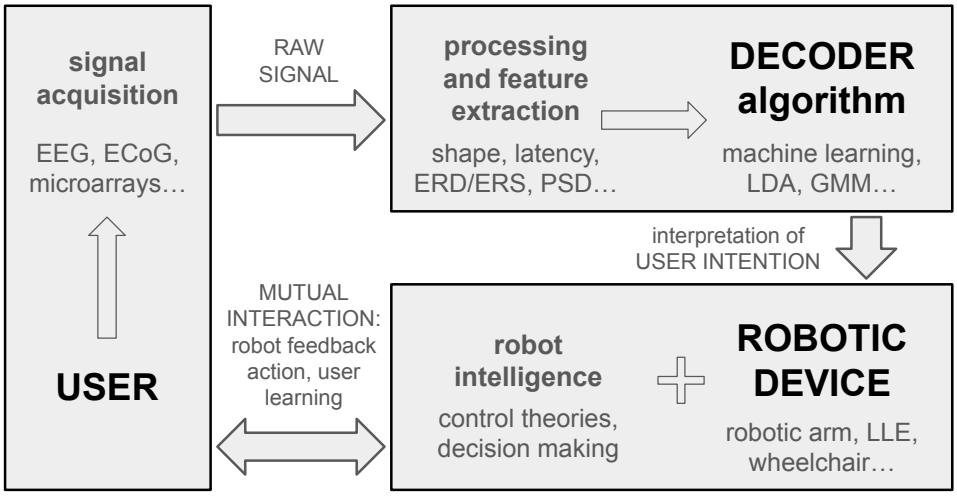


Figure 2.1: Typical generic BCI loop composed of three cooperative entities: user, adaptable decoder and cooperative device.

The lability of the interaction between user and decoders allows to chase the goal of improving the whole system performance through two main paths. Firstly, by focusing merely on the decoder: a posteriori training a bigger model or searching for alternative features to be considered or, in the case of deep learning networks, by training them with a larger quantity of newly acquired data. Otherwise, by focusing the attention toward the subject, and by allowing him to passively learn to produce better and more constant features, it is possible to perform acquisition and evaluation sessions according to the policy of 'infrequent decoder re-calibration' [20]. In a longitudinal study, with a total duration of 18 months, following the performances of a Cybathlon athlete, it has been shown that the user was "capable of significantly improving his performance over time even with infrequent decoder re-calibration" [21]. The duality of the calibration strategy allows to investigate a vast amount of training paradigms and to date a standardized, well-defined strategy guided by biomarkers is still unavailable.
Ultimately, to achieve performances capable of closing the gap between ideal obstacle-less environments and real world scenarios, the robotic device (which can be a prosthetic limb, a motorized wheelchair, vocal production device or any other actuator) has to evolve from a simple end effector to a cooperative machine capable of sensing the external world, planning the subsequent action, and performing it while considering in a collaborative manner the user intention [18].
As is imaginable, with such a wide definition, the applications of BCIs technology cover a vast number of needs, and depending on the features considered, there are many ways to classify the different typologies of interfaces.
Starting from a patient-centric vision of the three-entity complex, the first and most notable sub-

division of BCIs typology revolves around the brain activity acquisition methodology. To date, this essential step can be performed through invasive and noninvasive methods, both presenting radically different advantages and tradeoffs.

2.1.1 Invasive BCIs

To date the best achievements in terms of variety, accuracy and repeatability of results have been delivered by the utilization of BCI based on invasive cortical activity measurements [22]. Invasive recording technologies include micro-electrode array that capture individual neuron activity at the cortical level [22] and electrocorticography (ECoG) grids that record local field potentials in a less invasive manner directly from surface of the cortex [23] Although promising, these devices carry the main limitation of requiring a surgical operation to install the measurement electrodes array, therefore exposing the patients to potentially harmful interventions [24] that are today performed only on subjects already listed for neurosurgical session to treat pre-existing diseases. Furthermore, recent follow-up studies assessed the gradual formation of glial scars covering the implanted devices [25]; these scars represent a layer of self-defense of the cortex which aims at isolating healthy structures from wounded sections to prevent uncontrolled spread of damage. The formation of glia scars results in a cellular and biochemical barrier that greatly increases local impedance and affects the quality of electrical signals recorded by cortical electrodes, effectively compromising the decoding capabilities of the BCI system and implying the loss of the reacquired functions for the patient or the need for a renewed surgery [24], [26]. In addition, the implantation procedure itself is not exempt from risks such as infections, bleeding, or unintended damage to delicate cortical areas, which can permanently affect neurological functions. Even in cases where the implantation is successful, long-term stability remains a critical challenge: electrodes can shift microscopically over time due to natural brain movements, further reducing signal reliability [24], [26].

Another limitation is the difficulty of scaling these systems for widespread clinical adoption, since neurosurgical interventions are expensive, require highly specialized personnel, and may not be ethically justifiable in patients who do not already need brain surgery for medical reasons. On the other hand, the potential benefits of invasive BCIs are considerable, as they allow access to single-neuron or small-population activity with millisecond precision [27], which provides a level of control over external devices that is currently unobtainable with noninvasive approaches [28].

In the field of LLE control, invasive ECoG-mediated BCI solutions have been adopted to effectively translate upper limb motor intentions into stepping command achieving high and stable accuracies (84%) [29]. Another pilot study demonstrated the feasibility of using invasive ECoG signals to control a lower-limb exoskeleton in real time, showing that gait-related intentions

could be decoded and used to drive the device while also providing artificial leg sensory feedback [30].

2.1.2 Noninvasive BCIs

Noninvasive BCI technologies represent an import field of research to prevent the risks and complications associated to invasive procedures. The common characteristic of these interfaces is the use of external measurement instruments of cerebral activities. To date, recording neural activity from outside the scalp provides many challenges and numerous techniques have been developed to face most of them. Currently the technologies that are mainly adopted are functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), magnetoencephalography (MEG), and electroencephalography (EEG). As summoned in Figure 2.2 [31], each one of the previously mentioned technologies is characterized by its own advantages and drawbacks, and out of all of them, EEG technology is generally preferred because of its markedly better portability and temporal acuity.

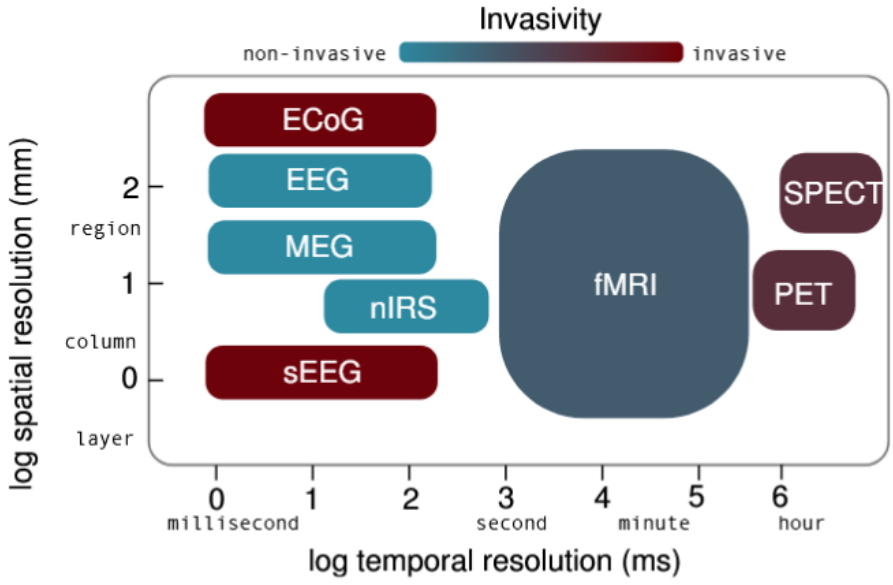


Figure 2.2: invasive and non invasive neuroimaging methodologies and their spatial and temporal resolution[32]

EEG acquisition

The acquisition of EEG measurements is mediated by electrodes positioned on the scalp of the subject. Those are kept in place by specific caps designed to locate electrodes at specific points described by international standards. To ensure the best possible signal acquisition and overcome the high impedance posed by skin-air contact, conductive gel application is often required;

this process is often regarded as the most limiting in view of a 'quick to setup' or 'self sufficient' BCI development. As expected, the high portability and low intrusiveness of EEG is paired by a notable difference in acquired signal quality: while through invasive electrode arrays it is possible to perform spike-sorting and recognize single cortical-neuron activations, EEG measures the cumulative postsynaptic electrical activity of populations of neurons, specifically the pyramidal ones located in the external layer of the cortex [33]. Due to their specific organization, they produce a perpendicular (in respect to the skull) current fluctuations detectable by the device.

It should also be mentioned that in between the pyramidal cell layer and the signal acquiring

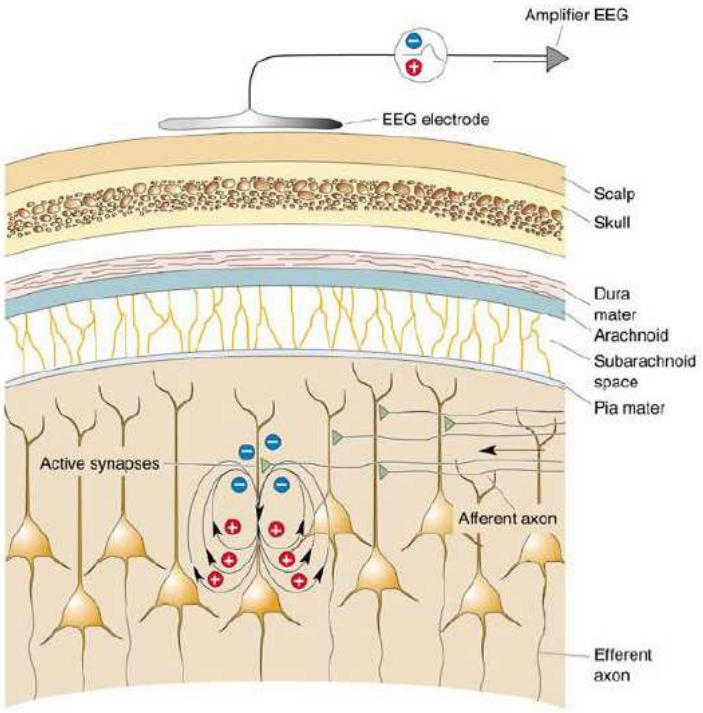


Figure 2.3: Pyramidal cell spatial orientation, postsynaptic potential and EEG electrodes [34]

device numerous isolating layers (such as skull, scalp, and cerebrospinal fluid) are interposed. The effect produced by those obstacles is comparable to that of a spatio-temporal low-pass filter that blurs the image and reduces the useful information to a bandwidth of 100 Hz. Furthermore, the brain is not the only electrically active organ in the human body; rather, among the skeletal muscles and the heart, it is the one producing the potential characterized by the lowest amplitude with recorded values that are typically below $150\mu V$ [33].

Nevertheless, thanks to frequency-based signal filtering, spatial sharpening, and other processing algorithms, EEG devices have been shown to be capable of producing discriminant features for task decoding.

2.1.3 BCIs paradigms

In the field of noninvasive BCI advancement, several paradigms have been developed to allow users to reliably instruct machines. Since the functioning of such systems depends on the user's ability to modulate cortical activity, the first classification of paradigms is typically based on the source of the influencing signal: changes in brain activity can be elicited by external stimulation (evoked potentials) or by endogenous voluntary modulation (spontaneous potentials).

Evoked BCIs

Evoked BCIs utilize the specific brain response to endogenous stimulation that can be rare (producing the so-called P300 wave) or erroneous (producing Error-Potentials). In this field, the most established paradigm is the so-called Steady-State Visual Evoked Potentials (or SSVEP). During SSVEP trials, the user is presented with a monitor displaying numerous flickering targets, each characterized by its ticking frequency. The instruction is to focus attention on the only target coding for the desired action. Over the years, it has been shown that the occipito-parietal activity of the brain, which encodes and processes retino-optical information, synchronizes with the activity of the target; therefore, by inspecting the frequency content of the EEG signal it is possible to instruct the system [35].
The previously mentioned P300 wave-based paradigm, also called odd-ball paradigm, revolves around the time domain analysis of the potentials evoked by infrequently presented stimuli in the central and central-posterior scalp positions [36], [37]. The wave name derives from its sign (Positive) and usual latency (300 ms after the stimuli presentation), although it is nowadays widely known that the actual time characterization varies as a function of the subject itself, its cognitive state, and the external environment influences. The detection of the P300 component revolves around the sequential presentation of many different stimuli from the whole range of possible commands (which may be of various nature, such as visual or auditory). The online applications function by analyzing the presence or absence of a P300 component in the measured response to each event, subsequently a vector representing the probability of each stimulus corresponding to the selected one is built, and only after a prefixed threshold is surpassed a decision is taken [38].
A similar approach exploits the brain's natural response to perceived mistakes. Error Potentials (ErrP) are event-related potentials typically recorded over frontocentral EEG sites. Their emerging occurs when a user observes an error, either in their own action or in the output of a system [40]. In this paradigm, the BCI system monitors the EEG alteration and evaluates the presence of these potentials to use them to automatically detect mistakes. Such systems can thus adapt or correct their behavior without explicit commands from the user, and have been utilized to reinforce the robustness of other BCI paradigms or to directly control spellers and robotic

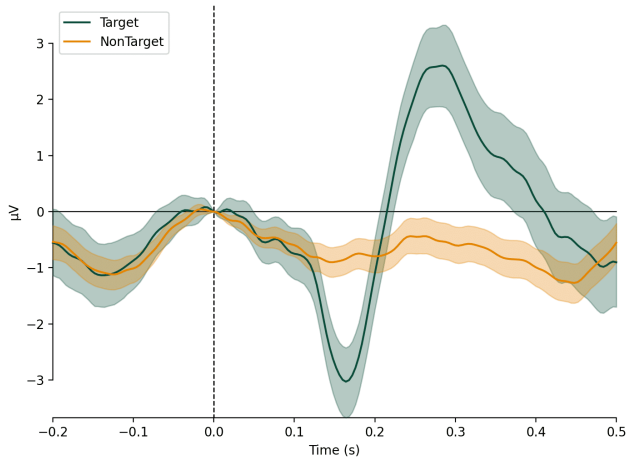


Figure 2.4: Average wave shape to target and non-target stimuli of a P300 paradigm [39]

devices [18], [41].

As can reasonably be inferred, although functional and accurate, these methods present many limitations mainly regarding the constant need of an external stimulus device and the need to code the user's intention into a different nature type of command. In particular, in the field of prosthesis command, as discussed before, the general aim is to enable the subject to interface itself with the device through a more natural type of command, such as thinking of moving the missing limb.

Spontaneous BCIs

To achieve a more natural form of control, spontaneous BCI paradigms have been developed. Among these, motor imagery (MI) represents the most common approach [18]. The rationale for decoding cerebral activity induced by imagined movements is well established and is rooted in the neuroanatomy of the cortex. It is now widely recognized that applying targeted electrical stimulation to multiple sites of the motor and somatosensory cortex elicits the corresponding responses in specific body muscles [42], [43]. Through such experiments, most notably those conducted by Penfield and colleagues [44] in the mid-20th century, the motor and somatosensory homunculi were mapped. These findings demonstrated that the brain encodes motor and sensory representations in a topologically organized fashion.

The second breakthrough that underpins the theory of motor imagery signal decoding concerns the characterization of brain rhythms and their correlation with cerebral activity. A widely accepted interpretation of these oscillations is that they reflect the intrinsic idle-state frequencies of cortical circuits. The most famous example is the bound between alpha waves (8-12 Hz) and

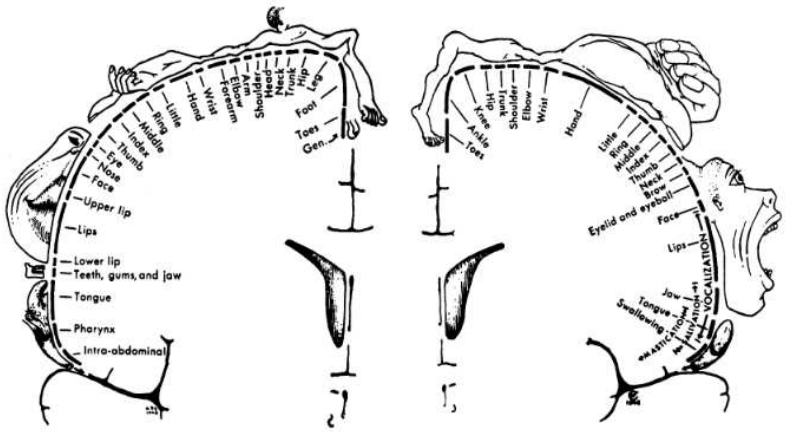


Figure 2.5: The motor and sensory homunculus. Penfield and Rasmussen, 1950 [44]

visual information processing networks located in the occipital cortex: as the main input source is shut down (corresponding to the closing of the eyes), the power density of the posteriorly localized signal of the alpha frequency band usually increases [45].
Similarly, for motor and sensory tasks, the location of interest is the sensorimotor cortex (localized underneath the frontocentra -FC-, central -C- and centralparietal -CP- row of electrodes in the 10-20 internationally standardized array) and the bands of main interest are the Beta band (13-30 Hz) [46] and the Mu band (8-13 Hz) [47], which is an anteriorly localized variant of the previously described alpha band.

Feet motor imagery and gait decoding

As briefly mentioned earlier, the vast majority of studies available in the literature focus on the analysis of hand or tongue motor imagery tasks [48]. By examining the homunculus in Figure 2.5, it can be observed that these body portions are commonly mapped far apart; furthermore, since they possess a much higher density of sensory receptors compared to other body regions, their cortical representation also occupies a considerably larger area of the brain [44].
Another consideration concerns the distinction between a simple foot, knee, or hip MI task and the complete gait imagination task. According to some publications, the frequency bands that carry the highest amount of useful information active during gait production tasks tend to shift toward slower rhythms, particularly within the [0.5-8] Hz range of the delta band [49]–[51].
Ultimately, Tortora et al. [52] in a recent study underscored the effect of LLE on cerebral modulation, concluding that ”exoskeleton activity does not modify brain patterns, but the constraint posed by the mechanical joints of LLE in itself (even if the user is actively performing the gait task) significantly modulate cerebral connectivity”, as a consequence it was considered essential to collect data and evaluate the system using an LLE device.

As a final remark on the topic of foot MI, it is worth emphasizing the broad consensus regarding the intrinsic challenges of distinguishing signals originating from the left and right foot MI. In general, foot-related motor imagery tends to generate weaker and less distinct EEG features compared to hand imagery [53], both in terms of spatial localization and spectral power. This attenuation is largely caused by the deeper and more medial cortical representation of the lower limbs, which results in signals that are less pronounced and more prone to overlap when recorded from the scalp [54], [55].
Moreover, despite the symmetrical positions, the proximity of the two feet representation leads to a minimal lateralization in the EEG patterns. As a result, distinguishing between the two is possible but becomes particularly difficult using non-invasive acquisition techniques[56].
Finally, it should be noted that this difficulty is paired by substantial inter-subject variability. Individual differences in motor imagery ability, cortical anatomy, and signal-to-noise ratio can markedly affect the reliability of EEG decoding, making the recognition of right versus left foot MI not only a technically demanding task but also one highly dependent on subject-specific factors.

2.1.4 BCI classification methodologies

To close the loop of any BCI system, it is necessary for the machine to interpret the user's intention and instruct the actuator. To cope with this task, real-time classification of EEG recordings has to be performed.
In early BCI development studies, the classification problem was often resolved using linear discrimination or Gaussian classifiers [57]. These methods, although successful, base their performance solely on the separability of space and frequency based features (as for the FS metric), but do not consider in any way the time-dependent nature of electroencephalographic signals. In this context, the need to model the temporal evolution of EEG signals and therefore handle an increasing number of features, encouraged the adoption of neural network architectures [18]. Artificial neural networks, firstly introduced by Warren S. McCulloch and Walter Pitts in 1943 [58], learn nonlinear transformations through layers of interconnected units, progressively extracting more abstract and informative representations [59]. The most basic unit is called perceptron, and its functioning schematized in Figure 2.6 is highly inspired to that of a neuronal cell. The unit performs a weighted sum of all its inputs and, after the application of a non-linear activation function (such as tanh, ReLu, or other), it passes the resulting value as an output to the next unit. The learning process consists of an iterative minimization of the discrepancy between predicted and target output. This operation is made possible by back-propagating the error through the network and updating the weights via gradient-descent–based optimization. After the learning process, such network can be utilized for classification and regression appli-

cation[59]. Such mechanisms allow neural networks to adapt to complex and high dimensional patterns, thus capturing dependencies that conventional linear methods cannot exploit.

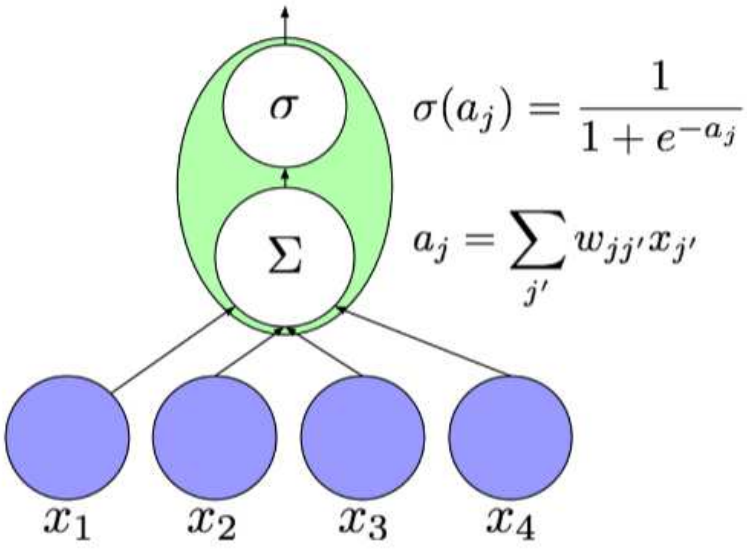


Figure 2.6: Basic perceptron structure: it computes a nonlinear function (σ) of the weighted sum (a) of its inputs ($x_{j'}$) [60]

To cope with the need of time-pattern consideration, Convolutional Neural Network (CNN) have been adapted to the analysis of EEG signals [61]. These architectures have been particularly effective in capturing spatial pattern across two-dimensional images and have been widely utilized for image recognition processes. Their functioning relies on the convolution operation through which small learnable sliding kernels (filters) detect local patterns such as edges, corners, or shapes [62]. Each kernel produces a feature map that can then be further elaborated. By stacking multiple convolutional layers, the network can progressively combine low-level representations into more abstract and meaningful features, enabling efficient and hierarchical learning. In EEG application, CNNs treats multichannel signals as two-dimensional spatial maps and kernels can be applied along the spatial or the temporal dimension in a distinct or common way. This approach has already proven to be effective for motor imagery classification tasks [61], [63]–[65].

In particular, in the domain of gait classification, a recently conducted study introduced the use of Recurrent Neural Networks (RNN) to address the lack of consideration of temporal dynamics by canonical classifiers [51]. As previously described, the problem faced during feet motor imagery or gait decoding classification is the overlapping nature of class distributions, which may otherwise differ if time patterns are considered.

Recurrent Neural Networks (RNN) have been specially designed to model sequential data. Unlike classical feedforward networks, which are developed to consider each input as independent from the previous one, RNN retains past information thanks to an internal state representation

[66]. This representation and recurrent modulation allows the contextualization of data points and is suitable to model the temporal dependencies that characterize EEG sequences. The most basic RNN architecture (Figure 2.7 is based on single repeating units that for every time step combine the information carried by the input vector (X_t) and the hidden state saved memory (H_{t-1}) to update its internal state (H_t).

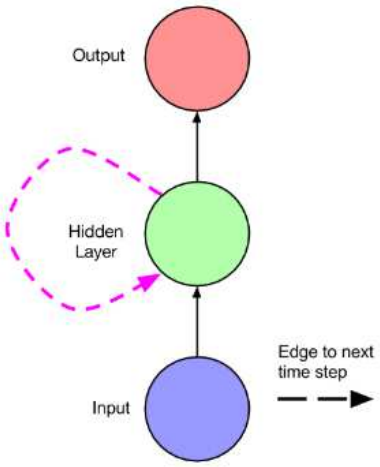


Figure 2.7: A simple recurrent net with one input unit, one output unit, and one recurrent hidden unit [60]

In recent years, more sophisticated recurrent neural architectures have been developed (such as GRU and LSTM) addressing issues such as the exponential explosion or vanishing of the gradient over time, and the challenges associated with short- and long-term memory retention [60].

2.1.5 Literature review

A concise literature review was conducted to position this work within existing research on decoding of lower extremities motor intentions or BCI-controlled LLE. Selected studies were compared in terms of neural features, decoding strategies, and exercise performed. The key findings are summarized in Table 2.1.

The first consideration concerns the distinction between motor decoding performed during actual gait production (which is typically aimed at identifying ongoing locomotor patterns) and motor prediction (which instead occurs prior to movement onset and is generally intended to enable proactive exoskeleton control). These two approaches differ not only in their temporal dynamics, but also in the type of neural information they rely on. Understanding this distinction is essential when comparing studies in the literature, as it influences both decoding strategies and performance. Studies that revolve around pure MI show a diffuse selection of μ- and β-centered frequency band [54], [56], [69], [76], while the studies aimed at gait decoding usually

consider also (or only) lower oscillations [49], [51], [63].
On a signal perspective, the majority of the publications is based on EEG acquisitions, but sometimes they are integrated with EMG or kinematic data to reduce noise sensibility or inter-subject variability. On the other hand, if the classification typology is considered, it is possible to affirm that classical algorithms such as Linear Discriminant Analysis (LDA), Support Vector Machines (SVM), and Nearest Neighbors (KNN) are the most utilized; nevertheless some deep learning based approaches are still observed [51], [63], [75].
The type of mental exercise presents huge variability, as MI exercises can be performed relatively to the movement imagination of single joints [54], [76], body portions [56], [69], or even relatively to more complex actions [50]. Finally, not all work include an exoskeleton-mediated feedback that as previously discussed may alter the results [52] or preclude the possibility of sustaining user-adaptation processes.

AUTHOR	EXO	ONLINE	TYPE	task	MEASURES	CLASSIFIER	f [Hz]	src
Tortora et al.	yes	no	decoding	gait	EEG EMG Kin	LSTM	[1,40]	[51]
Tortora et al.	yes	no	decoding	gait	EEG EMG Kin	LSTM	[1,8]	[49]
No-Sang Kwak et al.	yes	yes	predict	SSVEP	EEG	...	[...]	[67]
Sho Nakagome et al.	no	no	decoding	leg joint angles	EEG	RNN LDA	[0.1 50]	[64]
Hashimoto et al.	no	no	offline predict	feet MI	EEG	LDA	[13,30]	[56]
Ferrero et al.	yes	yes	predict	gait	EEG	LDA	[4 25]	[68]
Lee et al.	yes	yes	predict	hand MI	EEG	LDA	[14-19]	[69]
Blanco-Diaz et al.	yes	yes	decoding	gait	EEG EMG Kin	CNN	[0.1 4]	[63]
Kline et al.	no	no	decoding	gait	EEG EOG Kin	LDA	[α, β, γ]	[70]
Malik et al.	no	no	decoding	leg joint extension	EEG Kin	LDA	[8-30]	[71]
Bose et al.	no	yes	decoding	feet MI	EEG	KNN	[0-28]	[72]
Choi et al.	yes	yes	predict	stand/sit/blink	EEG EOG	SVM	[7-34]	[73]
Wang et al.	yes	yes	predict	step/stand/sit	EEG	SVM	[5 35]	[74]
Li et al.	yes	yes	predict	gait MI	EEG EMG	EEGNet	[4 40]	[75]
Yokoyama et al.	no	no	decoding	stance/swing obs	EEG	LDA	[0 50]	[50]
Tariq et al.	no	no	decoding	Knee MI	EEG	LDA	[7 32]	[76]
Tariq et al.	no	no	decoding	Feet MI	EEG	SVM LDA	[7 35]	[54]

Table 2.1

Chapter 3

Materials and methods

The system developed and evaluated in this thesis incorporates a feet motor-imagery BCI built upon deep neural network classifiers and delivers feedback to the user through both a lower-limb exoskeleton and a visual display. Its primary goal is to collect a comprehensive dataset across a three-session experimental protocol and to assess the feasibility of a rapid and accurate non-invasive control interface. The interface is designed to allow healthy participants to operate the mechanical device through movement intention to obtain a natural and intuitive control modality. The overall framework includes an EEG acquisition unit, a signal-processing and feature-extraction pipeline, a machine-learning classifier for intention decoding, and a control module that converts decoded commands into exoskeleton actuation. This section outlines the system's components and details the materials and methods used for data acquisition, processing, and validation.

3.1 System design

3.1.1 Materials

To set up the experimental environment, the following devices were employed:

- One master PC running the software for data acquisition, processing and classification. The monitor was used to guide the experimental protocol and provide feedback and instructions to the user through a video display.
- The open source lower-limb exoskeleton "ALICE" platform developed by the Interactive Technologies Studio INDI based in France [77]. It is a 4-Joint System focused on lower-limb support of patients with muscular dystrophy, cerebral palsy, spinal muscular atrophy, and other cerebellar disorders. To maintain the correct posture, participants utilized hand crutches.

- The 64 channel EEG acquisition cap waveguard by AntNeuro and its relative eego64 amplifier [78].

The connection between devices was mediated by cables to minimize the delays and ensure temporal synchronization.

Figure 3.1: ALICE exo [77], Antneuro waveguard cap and eego 64 amplifier [78]

In this work, 4 healthy participants (3 male and 1 female) were recruited, with an age in the range of 23 to 25 years old. No participants had previous experience with BCI systems or LLE-assisted gait exercises.
To identify them, an alphanumeric code was assigned to each participant: K9, L1, L2, L3.

Robot Operating System (ROS)

In order to integrate the different hardware and software components required for the experiment, the Robot Operating System (ROS) was employed as the development environment due to its flexibility in managing distributed processes and its extensive set of libraries [79]. Moreover, its open-source nature facilitates collaboration within the research community and allows the reuse of functional packages, thereby reducing the need to develop each component from

scratch for new projects. The framework currently supports several programming languages, including C++ and Python, and its full source code is publicly available.
The ROS structure revolves around packages. These are directories that contain an XML manifest that describes the package itself and specifies its dependencies. The core concepts underlying ROS include nodes, messages, topics, and services.
Nodes are single and independent processes that can perform any kind of programmable computation or algorithm. Their execution is independent of the execution of any other parallel nodes, so projects can be divided into different modules. This modularity enables the distribution of the tasks that the machine has to perform, thus promoting scalability and debugging.
Topics are the main communication channels for nodes. They are based on a publish/subscribe architecture: each node can publish a message on the topic while any other node receives it and performs any kind of callback action. This mechanism is therefore asynchronous and decoupled.
Messages are standardized data structures that are used to exchange information between nodes over a topic. Topics therefore accept only predetermined data types ranging from float numbers to strings, array, or more complex structures.
On the other hand, a synchronous communication mechanism is guaranteed by the utilization of Services that are based on a server/client paradigm. The client node sends a request to the server node, which processes it and provides a response.

ROS Neuro

As previously aforementioned, the open-source nature of ROS enables users to share working packages. Among those the publicly available ROS-Neuro posed as a perfect framework for BCI application thanks to prebuilt interfaces, acquisition, filters and other elaboration frameworks for EEG signals [80].

3.1.2 Protocols definition

To develop a working BCI the first step is to design valid, replicable and standardized calibration and evaluation protocols. Furthermore, since the vision behind the thesis is to help the user familiarize with new devices such as the BCI interface and the lower limb exoskeleton (de facto optimizing the time spent recording useful data with an EEG device), three additional protocols were implemented and tested during the acquisition sessions, raising the total number of protocols to five:

- Familiarization
- Calibration

- No-Exoskeleton Calibration
- Simplified Evaluation
- Evaluation

Calibration

During calibration, the subject's EEG activity is recorded while performing the desired imagination task. In particular, the subject was asked either to perform the imagination of a step (class *Movement*) or to relax (class *Rest*). To synchronize, randomize, and later access the data correctly, the user was instructed to follow the instructions presented by the monitor placed in front of him. To promote user learning, after each imagination task, the exoskeleton performed the correct movement as positive feedback. As can be observed in Figure 3.2 the ROS structure of the system is organized around four different nodes. /acquisition streams the raw acquired data from the EEG device that are stored, together with Event type messages into the .gdf file by the /recorder node. /rosserial is responsible for the control of the sensors an motors of the exoskeleton and transforms the movement command request into an actual step sequence. Finally, /smr_feedback is the central node that manages and paces all the processes.

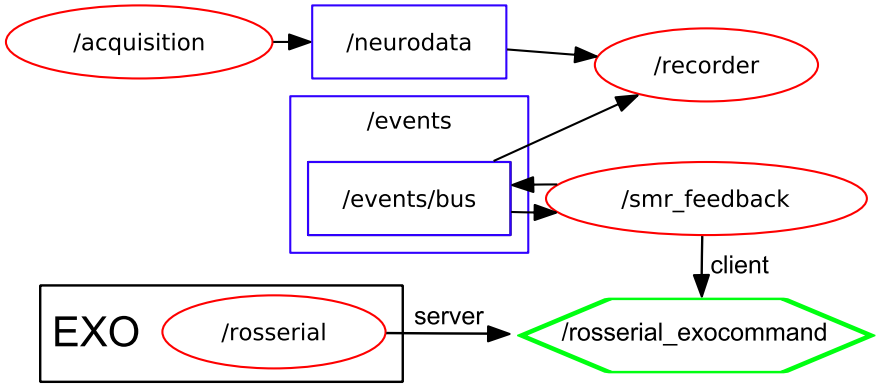


Figure 3.2: graphical representation of the ROS infrastructure of the calibration protocol. Nodes (red), topics (blue), services (green) and actions (arrows) such as publishing and subscribing messages on a topic or hosting and calling services.

A single calibration run was composed of a total of 16 trials, and each trial replicated the same sequences of visual instructions:

- a white cross (fixation cross) with a duration of 1 second which establishes the beginning of the trials and aims at resetting the user's attention.
- a colored dot (cue) paired with the corresponding task name and a visual representation of it. The cue remains on the screen for a total time of 2 seconds. The only three possible

cues are ‘left step’ (red dot), ‘right step’ (blue dot) and ‘rest’ (yellow dot). To aim at reproducing a natural movement sequence, the steps are always full-length steps and it alternates left and right instructions with randomly interposed rest periods.

- a green-colored dot establishes the beginning of the continuous feedback (CF) period, in which the subject should perform the imagination task or resting. It has a total duration of 4s for the movement tasks and 8s for the rest task. The dot is paired with a blue bar that fills automatically over time. At the end of the continuous feedback epoch the blue bar is completely filled and brightens itself, instructing the subject that the exoskeleton will start the movement corresponding to the trial’s cue.

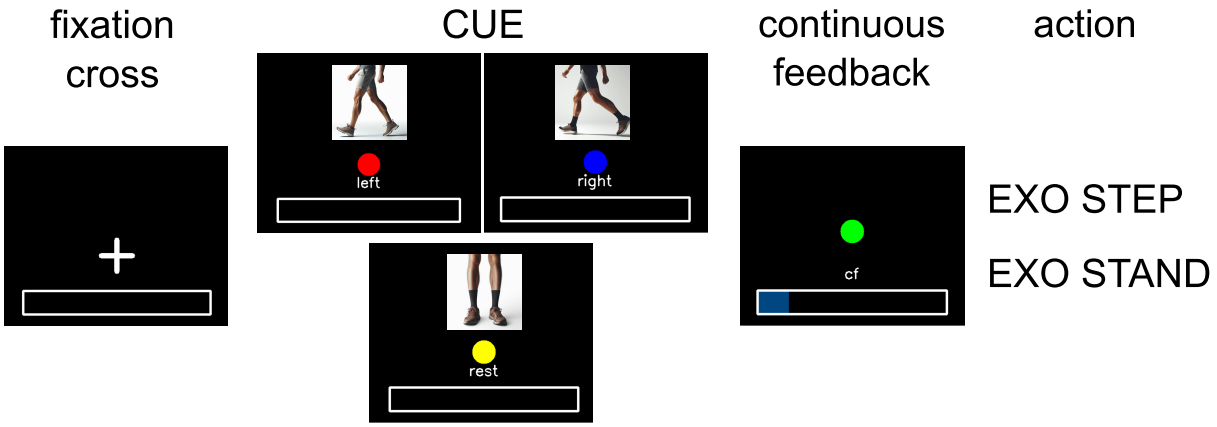


Figure 3.3: sequence of visual instruction provided by the monitor to the user, during calibration runs the blue bar fills automatically at constant pace over time.

No-Exoskeleton Calibration

The use of the LLE requires the support of crutches to maintain balance during walking movements. Consequently, the additional effort of the upper limb, required to support body weight with crutches, may contaminate participant’s EEG signals and hinder the initial identification of discriminative features necessary to distinguish rest from gait-related tasks. To minimize this effect, the proposed protocol also foresees calibration runs with the subject’s standing without wearing the exoskeleton. This protocol was used also to allow the user to familiarize with the BCI system before introducing the exoskeleton device.
To perform such acquisitions, a simplified version of the calibration protocol was developed. The only disparity from the standard calibration pipeline is the absence of the exoskeletal device and the ROS infrastructure required to manage it. In order to provide sensory feedback during runs, participants were instructed to self-perform the step described at the end of the continuous feedback period.

Familiarization

Since no subject had previous experience with the utilization of exoskeletal devices, to facilitate the learning process, a protocol discharged of any EEG-related node was designed as presented by Figure 3.4. As for the calibration runs, it is composed of repeating trials lacking the continuous feedback sequence since no signal is measured during this phase; it provides the participants the possibility to easily familiarize with the exoskeleton before its used with the EEG within the BCI system, to reduce the risk of recording contaminated trials.

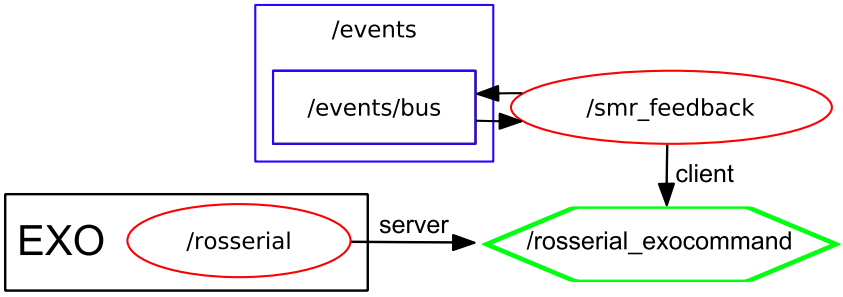


Figure 3.4: graphical representation of the ROS infrastructure of the familiarization protocol.

Evaluation

The protocol designed for online trials follows the same structure as the calibration sessions, consisting of alternating left–right cue sequences interrupted by randomly inserted resting trials. The main difference lies in the continuous feedback (CF) phase. In this case, the horizontal bar does not autonomously fill as during calibration; instead, it starts half-filled (corresponding to an uncertain state between Rest and Movement evidence) and moves according to the output of the integrated BCI outcomes. Its behavior is driven by the online classification of the EEG signal, processed in real time by a dedicated node. The node outputs the probability that the current EEG window represents a movement intention. This probability is then integrated over time: the bar moves to the right when a movement intention is detected, and to the left as the system identifies resting.

The control framework used to translate bar movements into exoskeleton commands follows a typical one-class BCI paradigm, in which the only active class corresponds to the "Movement" state. Within this framework, the rest condition is maintained as long as the decision threshold is not exceeded, while the feedback phase terminates as soon as the movement-intention threshold is reached.

The participant's task is to modulate their cognitive state according to the visual cues and control the bar accordingly. When the cue indicates a movement, the goal is to reach the predefined threshold as quickly as possible. In contrast, when the cue indicates rest, the participant must

avoid triggering a motor onset for the whole duration of the trial (10 seconds). In any scenario, the feedback period ends if the threshold is reached or if the time exceeds the timeout limit; subsequently the exoskeleton is instructed to perform the task that was presented by the visual cue.
In addition to the threshold value, the speed of the bar can be adjusted through a dedicated tuning parameter to refine the task dynamics.

Simplified Evaluation

A common issue in one-class BCIs is the classifier's inability to explicitly represent the resting state as a separate class. As a consequence, decision outputs often become biased toward a single outcome. To initiate the online testing phase with a more user-friendly evaluation protocol, as proposed in a previous study [81] a new simplified approach was implemented. The main idea behind this framework was to provide the participant with only positive feedback, thus avoiding any discouraging or misleading responses. To achieve this, the integrator node was modified so that the bar would move only in response to correct classifier decisions, thereby ensuring motion toward the appropriate direction.
The single-trial objectives were accordingly adjusted to reflect these modifications. For both tasks, the participant's goal was to move the horizontal bar towards the corresponding extremity (left for rest and right for movement) within 8 seconds. In this setup, any failed attempt had to be repeated once before proceeding to the next trial, and the exoskeleton was activated only upon correct movement detections.
This new strategy offers a twofold advantage. First, it helps participants familiarize with the evaluation protocol while gaining self-confidence through a consistent positive feedback mechanism; this aspect is known to play a key role in sustaining user learning and engagement [82]. Second, the modified integrator allows the inspection of possible classifier biases: since the bar does not move for incorrect class output, it becomes easier to identify and discard networks with strong class preference at an early stage.

3.1.3 Session design

Subjects were asked to undergo a total number of 3 experimental sessions. The proposed common framework consisted of:

- a first session focused on "No-Exoskeleton Calibration" modality aimed at getting used to the protocol and recording a data set large enough to search for meaningful features and train the initial classifier.

- a second session focused on the ”Calibration” modality with the exoskeleton, followed by ”Simplified evaluation” runs and eventually ”Evaluation” runs if acceptable results are achieved.

- a third and final session composed only of ”Evaluation” runs.

Each session lasted no more than two hours, considering the preparation of the setup, the user's instruction, and the experiment, to avoid the risk of user fatigue and habituation. All subjects performed 8 ”No-Exoskeleton Calibration” runs during the first session. On the other hand, in the following two sessions, the number of calibration and evaluation runs varied according to the specific subject performance. Nevertheless, to ensure the collection of a large enough evaluation dataset, at least 8 evaluation runs for each subject were recorded.

3.2 Calibration data analysis

3.2.1 Data acquisition

During calibration, the EEG data were acquired and recorded in a .gdf file. By analyzing the literature, it emerged a clear localization of foot-related brain activity in the central portion of the somatosensory cortex [44], [55], [56]. For this reason, out of the 64 available channels, only the most central 32 channels were considered. As highlighted in Figure 3.5, the location of the selected electrodes are chosen to be concentrated around the Cz channel. AFz is utilized as the ground electrode, and CPz is taken as the reference.
The EEG sample rate is set to 512 Hz, providing chunk of 32 samples with a frame rate of 16 Hz.
In addition to the EEG signal, time events related to the protocol phases and the exoskeleton movements are also stored in the recorded file to enable offline data analysis. In particular, events store information about their typology, starting time point, and total duration. The code numbers paired with the relevant instants are summarized in the following table.

786	fixation cross onset
700	rest cue onset
702	right step cue onset
704	left step cue onset
781	continuous feedback onset
898	hit
899	miss
333	movement onset

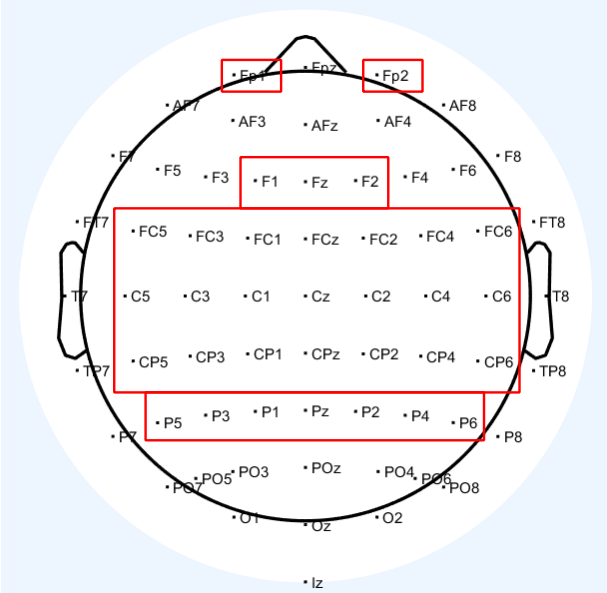


Figure 3.5: Antneuro 64 channel distribution the 32 channels highlighted are the one selected for the experiment. CPz is used as reference.

3.2.2 Preprocessing

EEG signals can carry a lot of information, but are known to be highly affected by noise and external artifacts. The source of signal distortion can arise from environment-related sources or from the body itself [33]. The main types of physiological artifacts comprehend:

- Ocular artifacts, caused by movements and blinks of the eye. This typology of signal pollution is directly related to muscle potentials and is characterized by a distinct frontal and prefrontal localization and a low frequency component.

- Muscle artifacts related to facial muscle activity, but also neck movements and swallowing. The frequency spectrum is often very large with components ranging from 20 up to 200 Hz and a diffused localization, although they are mainly recognizable in periferic locations.

- Cardiac artifacts, related to heart pumping activity. They can be observed mainly on electrodes placed near the main blood vessels and are not primarily related to the electrical activity of the heart, but rather to the mechanical effects of arterial expansion and relaxation induced by blood pressure. These pulsations may cause small displacements of the electrodes, resulting in periodic fluctuations in the recorded signal.

- respiratory artifacts, caused by thoracic movements. They present an extremely low-frequency component.

On the other hand, artifacts can also be produced by:

- Line Grid interferences. These are the most noticeable ones because of the evident peak at 50 Hz in Europe (60 Hz in USA) and other multipliers (100, 150, 200 Hz …)
- Movements of body segments or cables which may cause a physical dislocation of electrodes. In the time domain they can be recognized as unusual and out-of-scale spikes on one or more channels (simultaneously in this case)
- Equipment such as the amplifier, computer alimentation or exoskeleton actuators interference

To minimize the effect of these types of unwanted signal sources and improve the signal to noise ratio a series of preprocessing steps were performed.

Frequency filters The first and most essential preprocessing step consists in attenuating undesired low-frequency drifts and high-frequency noise components. For this purpose, a 5th order Butterworth bandpass filter with cut-off frequencies set at 2 and 40 Hz was employed.
Butterworth filters belong to the class of Infinite Impulse Response (IIR) filters and are designed to provide a maximally flat amplitude response within the passband, followed by a smooth, monotonic roll-off outside the selected frequency range [83]. The steepness of the transition band increases with the filter order: higher orders yield sharper frequency selectivity but also introduce a greater phase delay in the filtered signal.
The adopted cut-off frequencies were chosen to selectively remove slow trends caused by electrodermal, cardiac, or respiratory artifacts, while simultaneously attenuating high-frequency components such as line noise, muscle artifacts, and other noncortical sources; in this way this stage ensures that only the most informative frequency content, relatively to motor imagery, is retained for subsequent analysis.

Spatial filters In addition to frequency filtering, spatial filtering was applied to further enhance the signal-to-noise ratio by reducing spatially correlated noise and improving the spatial resolution of the EEG recordings.
Specifically, a Laplacian spatial filter was adopted. This method recomputes the signal at each electrode as the difference between its potential and the weighted average potential of its surrounding neighbors.
Mathematically, this operation acts as a spatial two-dimensional high-pass filter, emphasizing the local cortical activity beneath each electrode while suppressing the common background activity shared between adjacent sensors. The Laplacian filter is particularly effective in highlight-

ing activity originating from localized cortical regions, such as the sensorimotor areas typically involved in motor imagery tasks.

3.2.3 Feature extraction

The output of the proposed pipeline still represents a highly dimensional datatype that is difficult to interpret. To enable human inspection, more complex mathematical manipulations were applied to produce informative scores or more detailed representations of the same dataset. During the following chapters, a single (EEG-channel; frequency-band) pair will be referred to as a feature.

Log-Band Power and Event Related Desynchronization

Since motor imagery task aims at promoting the modulation of the power in the sensorimotor rhythms, the first analysis was focused on the Log-Band Power (LBP) computation. This quantity measures on a logarithmic scale the relative powers that a specifically selected frequency band carries on a determined time window. In this study the analyzed bands are δ ([1-8]Hz), μ ([8-13]Hz)and β ([13-32]Hz).
Since the aim of the preprocessing step is to evaluate if (and how) this measure varies over time, it is essential to compute the evolution of the LBP over the course of the continuous feedback period. To the cause complex mathematical methodologies such as the Welch method have been developed, but a fast and computationally light approximation in the time domain was preferred. The pipeline starts with the time-domain filtering of the selected bands with a 5th order Butterworth bandpass filter, followed by a squaring operator and a moving average time filter. Based on the results of a previous study [84], a window length of one second was adopted. Finally, a logarithmic transformation is applied to reduce the variance.

$$LBP(t) = log(\hat{P}(t)) \tag{3.1}$$

$$\hat{P}(t) = \frac{1}{T} \int_{t-\frac{T}{2}}^{t+\frac{T}{2}} x(\tau)^2 d\tau \tag{3.2}$$

To visually inspect the data, it is helpful to assess the variation of the LBP relative to the power of a reference period. In this context, the baseline period was identified as the time window that includes the fixation cross and the presentation of the visual cue. This measure is called Event Related Desynchronization / Synchronization (ERD/ERS) and is computed as $ERD_{\%}(t) = 100 * (\frac{A(t)-R}{R})$. During a motor imagery task, a desynchronization of neuronal populations involved in motor processing is expected [85]: while at rest these circuits exhibit

rhythmic idling activity, their involvement in motor computations leads to a reduction in synchrony, reflected as an ERD. This phenomenon typically appears as a negative modulation of the power carried by specific frequency bands over the cortical regions associated with the imagined movement.

Fisher Score Maps

A statistical tool helpful in assessing the separability of different classes in a feature distribution space is the Fisher Score (FS). To provide a measure of how well the analyzed feature can discriminate between classes, it takes into account both the mean and the variance of the distributions that are assumed to follow a Gaussian curve. The mathematical formulation of the FS is as follows:

$$FS(x) = \frac{|\mu_1 - \mu_2|^2}{\sqrt{\sigma_1^2 - \sigma_2^2}} \tag{3.3}$$

Here μ and σ denote respectively the mean and variance of the i-th class distribution. By definition, the range of values that the score can possibly provide as output ranges from 0 to infinite. A potentially informative feature for a linear class classification problem is paired with a high FS, which indeed indicates a relatively high mean distance compared to the cumulative intra-class variability. On the other hand, low scores suggest that the analyzed feature will not be discriminative for the task. In the context of EEG-based BCI development, the described score is usually computed across electrode and frequency band pairs providing detailed spatial-spectral maps representing the feature relevance for the task of discriminating samples of EEG signal recorded either during rest or imagined movement. In our application, the feature maps are plotted as topoplots of the scalp for every frequency range analyzed.
The value of these maps is appreciable during the feature selection phase thanks to the easily interpretable information provided; but also to assess the potentiality of a single run to be meaningful as a training or validation dataset sample.

Power Spectral Density

The previously described feature computation methods allow for a precise spatial and temporal analysis of how the EEG signal varies over time, but reduce the acuity in the frequency domain. Through the Welch's method, it is possible to evaluate the evolution of the power spectral density (PSD) with higher frequency resolution.
The Welch's method is a refined spectral estimation approach that reduces the high variance typical of the standard periodogram. It operates by segmenting the EEG signal into partially overlapping windows. Each segment is multiplied by a tapering window function to reduce spectral leakage, and a periodogram is computed for each window. These segment-wise peri-

odograms ($P_i(f) = \frac{1}{N}|X_i(f)|^2, X(f) = Fourier[x(t)]$) are then averaged to obtain a smoother and more reliable PSD estimate across the analysis interval, making it particularly suitable for tracking changes in rhythmic brain activity across specific frequency bands.
Finally, to inspect possible differences between tasks, ERD/ERS have been computed using the rest task as a reference signal.

3.3 Classification

To close the loop of the BCI system, it is necessary for the machine to interpret the user's intention. To cope with this task, real-time classification of EEG recordings is performed. In this process, signal samples of a predefined temporal length (1 second) are continuously fed to a classifier at a fixed frequency of 16 Hz. The classifier's output corresponds to the predicted class label assigned to the sample, together with the probability associated with the classification. To obtain a more natural and reliable BCI outcome, an evidence accumulation framework has been developed to temporally integrate the constant output stream of the classifier. The ROS

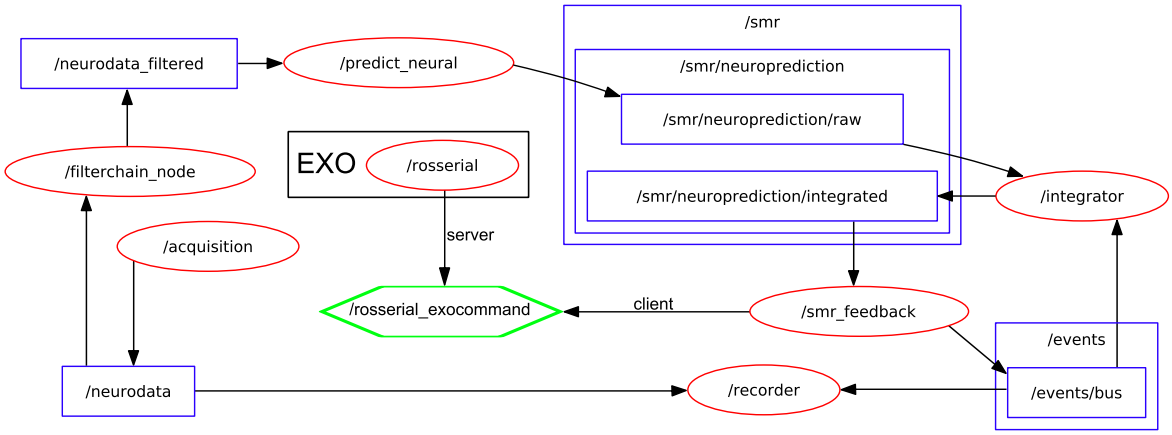


Figure 3.6: graphical representation of the ROS infrastructure of the evaluation protocol. Smr_feedback is the central node that manages and paces all the processes. Predict_neural is the node that requires the classifier architecture and weights and performs classification.

infrastructure that deals with the task described above is modeled in Figure 3.6. As it is possible to observe, with respect to the set of processes that enables the calibration runs, three additional nodes are activated, namely: /filterchain_node, /predict_neural and /integrator. The first independent process (/filterchain_node) subscribes to the /neurodata topic, on which the acquisition node publishes data samples of 512*32 data points (frames * channels) and applies the spatial and frequency filters. Subsequently, the /predict_neural node passes the filtered values through the fixed and pre-trained weights of the classifiers, and it provides a /neuroprediction output at a frequency of 16 Hz. This datatype carries hard- (class belonging) and soft- (probability

of class belonging) classification information. To the same topic subscribes the second node which, through different modalities, can perform the time integration of the classifier decisions. The decision output of the node is then published on the /Integrator/Neuropred topic, and it is the final value used by the central /smr_feedback node to control the displayed bar.

3.3.1 Classifiers

Since the aim of the current study is to provide an early and optimizable framework for LLE-BCI development, the design tries to minimize the variability of the protocol, thus a limited number of classifiers was employed. Based on the literature review, three EEG-adapted deep learning networks were selected to be trained and tested on the calibration data set of each subject. For the online evaluation, only the best performing networks were deployed with the possibility of performing a total recalibration or finetuning when necessary.

EEGNet

The first selected network is EEGNet, the first convolutional neural network introduced in the field of EEG-based BCI [61]. Developed and released in 2018, it is a highly generalizable and compact architecture designed to perform effectively even with limited training data. EEGNet is suitable for a wide range of BCI paradigms.
Its structure simplified in Figure 3.7 relies on three distinct convolutional strategies aimed at capturing meaningful patterns over different dimensions of the EEG signal.
First, temporal convolutions act as frequency-domain filters, enabling the network to learn frequency-specific representations.
The spatial convolutions then optimize frequency-dependent spatial filters across the EEG channels.
Finally, a novel separable convolution layer integrates temporal summaries of the extracted feature maps and linearly combines them across spatial dimensions, yielding a compact yet expressive representation.

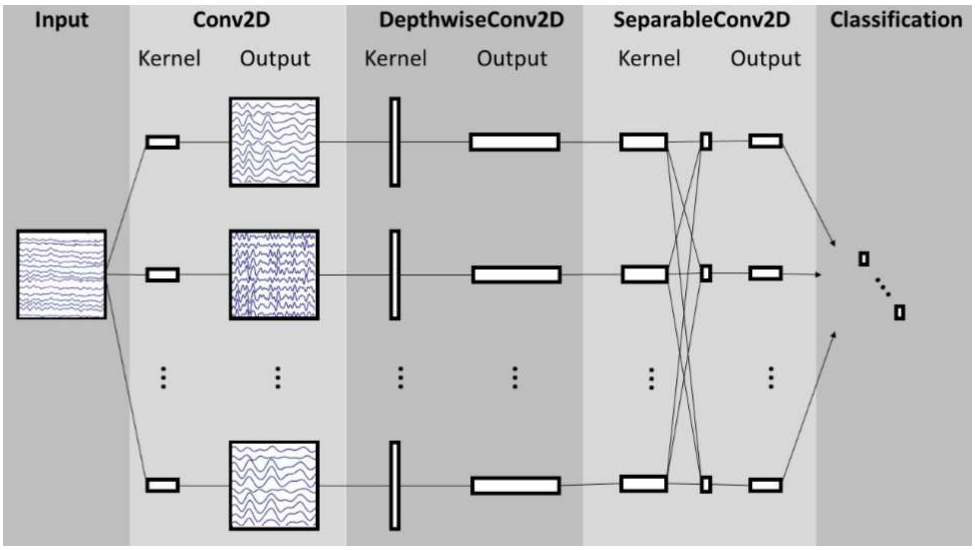


Figure 3.7: EEGNet simplified internal structure [61]

Thanks to it's architecture EEGNet offers also the possibility to perform an inspection of the trained convolutional kernels, thus promoting human interpretation of the results.

TGCN

A more advanced convolutional neural network selected for the same task is the so-called Temporal Graph Convolutional Network (TGCN) [86]. This class of models also relies on convolutional-style modules to detect discriminative patterns, but unlike the architecture of EEGNet, which treats EEG data as a standard multichannel time series, TGCNs explicitly account for both spatial and temporal structure simultaneously.

In particular, the model views multi-channel EEG as a structural time series: each electrode channel is treated as a node in a graph whose edges reflect spatial relations (e.g. scalp topology or functional connectivity) and whose node attributes evolve over time. The convolution operations are thus applied across both spatial and temporal dimensions: filters are localized and shared across time-steps and across node neighborhoods in the graph, yielding spatio-temporal feature extraction.

This is analogous to performing a 2D convolution on an image (where one axis is space and the other time), rather than applying two separate 1D convolutions in sequence.

In the application to seizure detection, the TGCN achieves two key advantages: first, the built-in invariance to when and where a pattern occurs supports generalization across variable electrode setups and temporal offsets; second, the inductive bias introduced by shared spatio-temporal filters allows the model to remain relatively parameter-efficient, which is especially valuable when training data are limited [86].

Both considerations are directly relevant to the field of motor imagery-based BCIs. Similarly to

seizure detection, the task involves identifying transient neural activations that deviate from a baseline or resting state (movement–seizure vs. rest–baseline). In this context, the hypothesis is that TGCNs can effectively model the spatio-temporal dependencies that characterize the transition between rest and movement-related cortical patterns. Moreover, the parameter efficiency of the TGCN architecture makes it particularly suitable for MI applications, where the available training data are typically limited due to subject-specific variability and time-consuming data collection procedures.

PhiNet

The final neural architecture considered in this work is PhiNet, a recently introduced model (2024) developed by a team of Japanese researchers [87]. PhiNet aims to emulate the internal dynamics of temporal prediction processes that are hypothesized to occur in the hippocampal formation.
As discussed by Chen et al. (2024) [88], the hippocampus can be modeled as a predictive autoencoder receiving input from the entorhinal cortex (EC), which is known to store spatio-temporal information. The proposed mechanism relies on the comparison of a true sensory observation with a delayed prediction generated from past activity. In this model, the CA1 region functions as an error-computing unit that receives two inputs:

- a direct pathway from the EC, exhibiting a short transmission delay (≈2.5 ms)
- an indirect pathway that passes through the dentate gyrus (DG) and CA3, introducing a longer delay (≈9–17 ms) due to at least three successive synaptic transmissions.

The CA3 subfield is hypothesized to implement a recurrent predictive mechanism that attempts to forecast the next EC input. The discrepancy between prediction and actual observation, computed in CA1, serves to iteratively refine the predictive model, thus forming what is described as a self-supervised cortico-hippocampal loop.

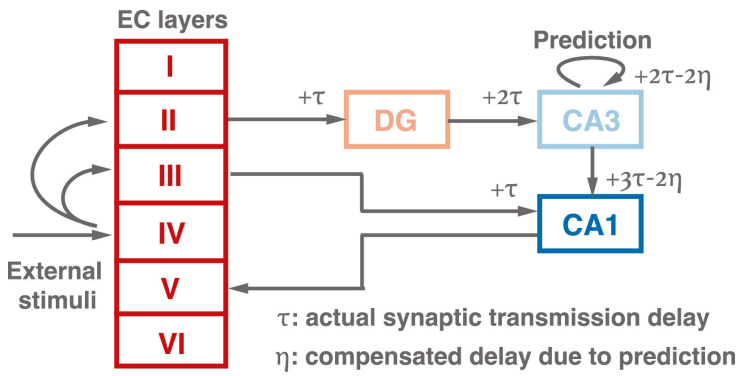


Figure 3.8: Model of hippocampal circuits [88]

Building upon this biological hypothesis, PhiNet implements a non-contrastive self-supervised learning framework inspired by the same predictive principle. The model generates two augmented representations of the same input and learns to predict the more complex representation from the simpler one, effectively minimizing the temporal prediction error.
As illustrated in Figure 3.9, the network architecture mirrors the organization of hippocampal formation: predictions generated by the CA3-like module are propagated to the CA1-like comparator, where a loss function quantifies their mismatch and back-propagates the error to refine the predictive units. Three deep encoders emulate neocortical (NC) processing pathways, and an explicit long-term memory component is introduced to prevent information loss and stabilize temporal representations. The inclusion of this network in the present study enables a com-

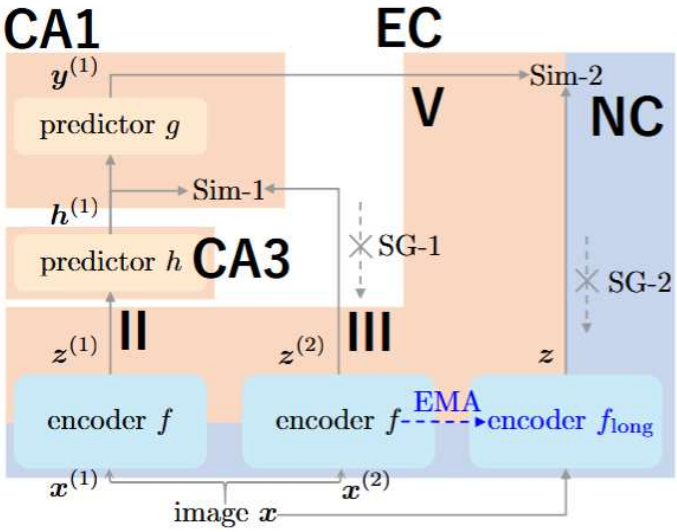


Figure 3.9: Simplified PhiNet architecture representation [87]

parison between the performance of conventional CNNs and that of a biologically plausible architecture, whose application to real EEG neural data remains partially unexplored.

3.3.2 Evidence integrator

The outputs of the classifiers node describe only a single time-point prediction of the participant's current neural state. As such, these measures are highly affected by noise, and single-sample accuracy rarely achieves satisfactory results.
To improve BCI performance, the temporal dimension can be exploited: by considering the evolution of classifier evidence over time, it becomes possible to integrate information across successive samples and capitalize on the classifier's typically small advantage over random choice.
Several strategies have been proposed to achieve this, ranging from simple majority voting

schemes (which merely count discrete class predictions within a temporal window) to complex dynamical-system-inspired models that integrate classifier confidence using multiple tuned parameters [89]. Given the limited number of recording sessions available in this study, and the consequent inability to perform extensive hyperparameter optimization, an intermediate approach was selected.
The chosen integrator initializes a tuple of n values (corresponding to the number of classes), each starting from the random choice probability (e.g., [0.5, 0.5] in our binary scenario). At each time step, the confidence of the classifier for the i-th class, denoted as $C_i(t)$ minus the random choice probability p_{rnd}, is added to the accumulated value $v_i(t-1)$, scaled by a parameter ev:

$$v_i(t) = v_i(t-1) + \frac{C_i(t) - p_{rnd}}{ev} \tag{3.4}$$

The values are clipped to remain within the [0, 1] interval.
In this two-class configuration, the values v_{Rest} and v_{Step} remain symmetric around 0.5. Since the 'Step' class is the only one associated with an active command, the feedback bar presented to the user displays only v_{Step}. As previously introduced, the proposed integrator depends on two main parameters:

- The decision threshold th. It ranges between 0.5 and 1 and defines the sensitivity of the system to the Step class. Its tuning depends on prior knowledge of the bias of the classifier obtained from simplified evaluation runs. Classifiers biased toward Rest should use a lower threshold (0.65–0.75), since maintaining a stable movement-related cognitive state can be challenging for the user. In contrast, classifiers that are biased towards movement, risk triggering unwanted false activations; therefore, a higher confidence requirement (threshold >0.85) is preferable.

- The ev parameter. It acts as a decision-velocity regulator: high values slow down evidence accumulation, requiring more consistent classifier outputs to move the feedback bar by the same amount. Its regulation highly affects the reactivity of the system; nevertheless it is often useful to slow down the decision process to allow the user to have more time to react to unwanted initial outcomes.

3.3.3 Evaluation metrics

Different measurements have been considered to evaluate the performance of the proposed BCI system. The classic measure is the single trial accuracy, but a more comprehensive statistics which balances performance both in terms of precision and recall is the F1 score.

$$\text{precision} = \frac{TP}{TP + FP}$$
$$\text{recall} = \frac{TP}{TP + FN}$$
$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN}$$
$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

Where T and F stand for True and False, while P and N stand for Positive and Negative. In this application the true class is associated with the movement.
Another important aspect to consider is the time required by the system to detect a movement intention. This time must be balanced with the capability of the system to avoid the sending of false positive commands during the rest trials. To give a quantitative evaluation of the system's reactivity and robustness, a normalized time score metric is introduced as:

$$normTimeScore = \frac{(timeout - \hat{T}_{Rest}) * n_{rest} + \hat{T}_{Mov} * n_{Mov}}{timeout * 2 * n_{tot}} \tag{3.5}$$

The timeout parameter is fixed in our protocol at 10 seconds and measures the maximum duration of the continuous feedback period. An ideal BCI system should have a normalized time score as close as possible to 0, indicative of high reactivity to movement intention and a low incidence of misclassified rest tasks. Additionally to trial-based performance metric, I also considered single-sample accuracy and F1 score to evaluate the classification performance independently from the integrator parameters.
For the requirements of the protocol, a simple two-class problem, a random classifier is expected to perform with an accuracy and F1 score of 0.50. With a sufficiently large dataset, any classification yielding scores >0.5 would be accepted as a better option. This is not the case for the datasets acquired during the trial, as it counts a finite number of observations.
In this regard to affirm that a selected decoder performs statistically better than a random choice, it is essential to perform a unilateral binomial test: a statistical test used to determine whether the proportion of 'successes' in a series of n independent Bernoulli trials (each with two outcomes: success/failure) is significantly greater or smaller than a hypothesized probability $p0$. It is possible to evaluate the p-value associated to the probability of observing a number of successes at least as extreme as $p0$ in the direction of the alternative hypothesis ($H1 : p > p0$). By labeling the number of total observations as n and the number of observed successes as x, p is calculated

as:

$$p\text{-value} = \sum_{k=x}^{n} \binom{n}{k} p_0^k (1-p_0)^{n-k} \tag{3.6}$$

The number of observations of single-trial accuracy measures corresponds to the number of trials within a single run (i.e., 16 trials), thus the minimum accuracy required to surpass the 0.05 confidence interval is 75%, while if the performances are evaluated on a 4-run basis (that corresponds to the minimum number of runs of the last session), the threshold drops to 59.4%. For single-sample accuracy, the number of observations varies from trial to trial but is in any case higher than 4000 for each of the online files, therefore the minimum accuracy for a significant classifier corresponds to 51.3%.

Chapter 4

Experimetral results and discussion

4.1 Features analysis

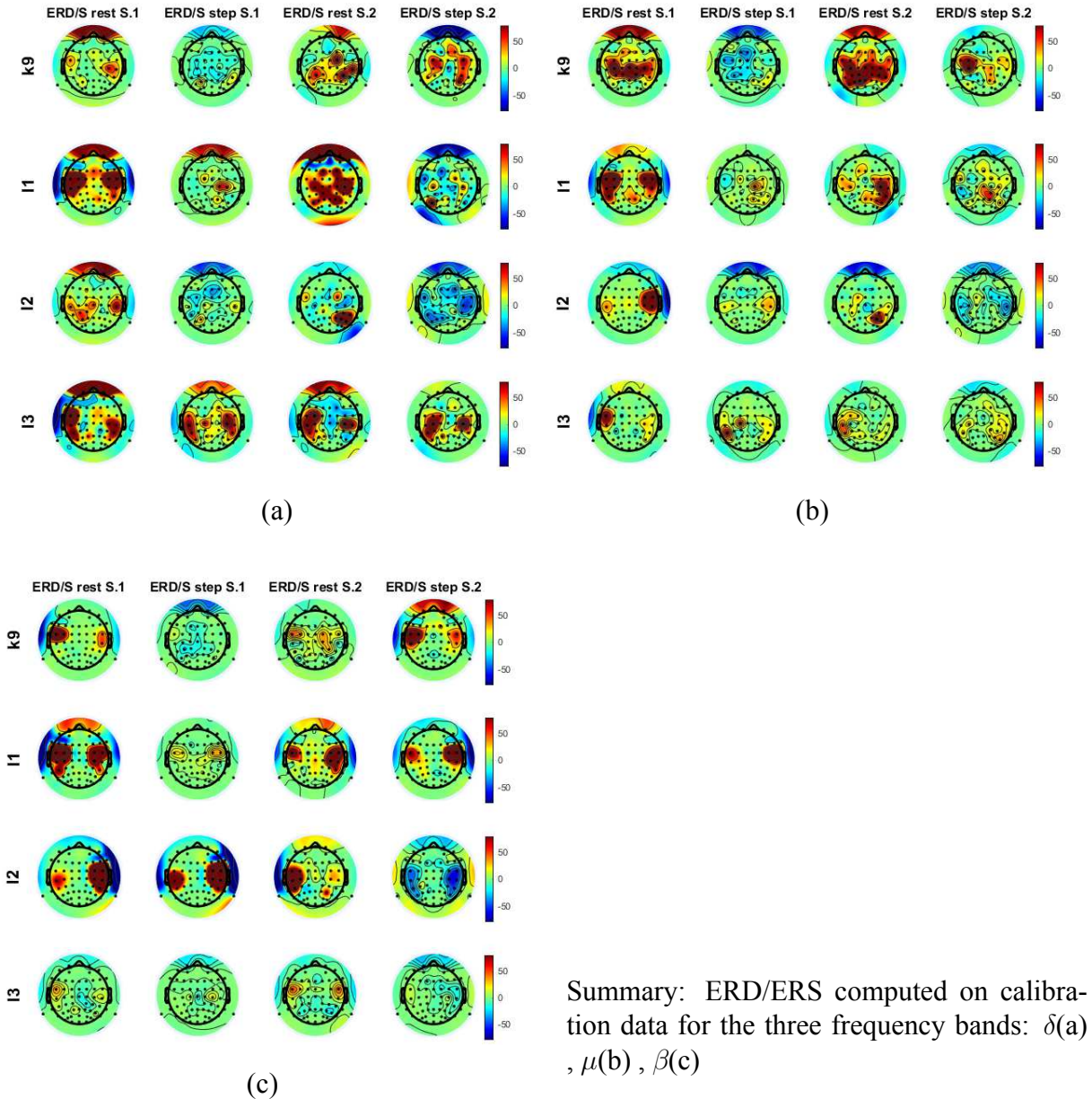
4.1.1 ERD/ERS

Figure 4.1 allows the inspection of ERD/ERS plots. The comparison of activity differences between the two tasks confirms, for the majority of the cohort, the expected theory-driven hypothesis: during foot motor imagery, the user mental activity tends to desynchronize on central electrodes at least in one of the three selected frequency ranges. Participant K9 is the one that exhibits the most appreciable results during the first session with high desynchronization values at the F1, Fz, FCz, Cz, CP1, and CP2 locations and throughout the whole analyzed spectrum. The peaks are reached by μ motor rhythms during the first round of acquisitions. In the second session, a clear shift of movement-related ERD toward the Cz and FCz electrodes is observed, mainly within the δ range.

For subjects L1 and L2, the highest desynchronization appears to occur in the lowest portion of the spectrum, δ, reinforcing evidence of its correlation to gait related movement planning [52], with the signal of interest generated by the sensors located over Cz and in the frontal region. Furthermore, both users show a shared increase in brain power (over all the bands) during rest tasks, localized on the extremities of the central coronal midline on sites which have been widely associated with hand activity. This activity peak can be related to the utilization of forearm crutches as assistive devices during the recording of the first and second sessions, and it is therefore not unexpected to detect it also during task-related periods.

Subject L3, on the other hand, is the only participant who does not show the classical expected outcomes: it appears that both task executions are associated with a centrally localized ERD in the β and δ bands, and the simple visual inspection of the topoplots does not seem sufficiently informative to establish whether quantifiable differences exist between the two conditions. By

inspecting the same user's plots, it is also interesting to observe the high synchronization that arises during both tasks equally in the delta frequency range on the electrodes localized onto the hand motor-controlling portion of the sensorimotor cortex. As discussed previously, this outcome is likely a consequence of a strong reliance on hand-mediated use of crunches to maintain balance, and it is likely that such high-intensity noise source may have overshadowed the activity changes caused by the execution of the mental exercise.



Summary: ERD/ERS computed on calibration data for the three frequency bands: δ(a) , μ(b) , β(c)

Figure 4.1: Summary of ERD/ERS patterns across the three frequency bands.

4.1.2 Fisher Scores

In this work, I present an initial analysis of the temporal evolution of FS maps across the calibration files collected during the early stages of the experiment as this are the one available during the process of BCI development. Specifically, for participants K9, L1, and L2, a total of 10 calibration files were acquired, with 8 collected during the first session and 2 during the evaluation sessions. In contrast, due to technical issues, participant L3 provided 6 files in the first session and other 3 were acquired in a subsequent second calibration session (1b). The interval in days between two consecutive sessions is reported by Figure 4.2; for participant L3, an additional 3-day interval occurred between sessions 1 and 1b. FS maps were computed for

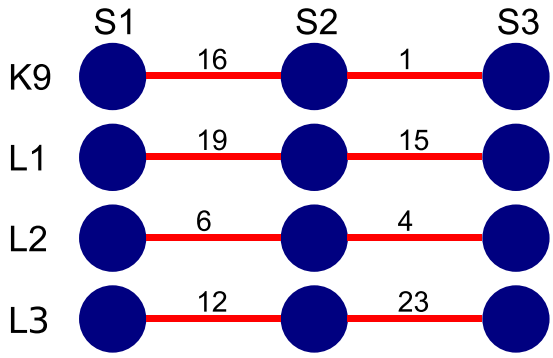


Figure 4.2: interval (in days) between two consecutive sessions

each file, allowing a visualization of how discriminative neural patterns emerged or changed over time. To further evaluate the stability or evolution of these patterns, the same metric was recomputed at the end of the full experimental protocol using the last four evaluation files of each subject.

Calibration As for ERD/ERS analysis, by analyzing the FS maps of calibration data illustrated in Figure 4.3 4.4 4.5 for each individual participant, it is possible to identify a common pattern of centralization at least in one frequency band for each user. The detection of the described phenomenon is more appreciable in the lowest frequency domain (Figure 4.3) since it is shared by all participants with lowest magnitude for K9. The Cz, Pz and FCz electrodes are the sites toward which the most discriminat features appear to converge over time.
The sensorimotor rhythm μ tends to exhibit a peculiar run-to-run consistency, and the corresponding Figure 4.4 illustrates the distinct stability of the Fisher Score in the fronto-central portion of the cortex, a feature shared by all subjects. The second most prominent region is the central-posterior area, which, in comparison, tends to manifest a lower consistency across runs and reaches high scores only in sparse runs.

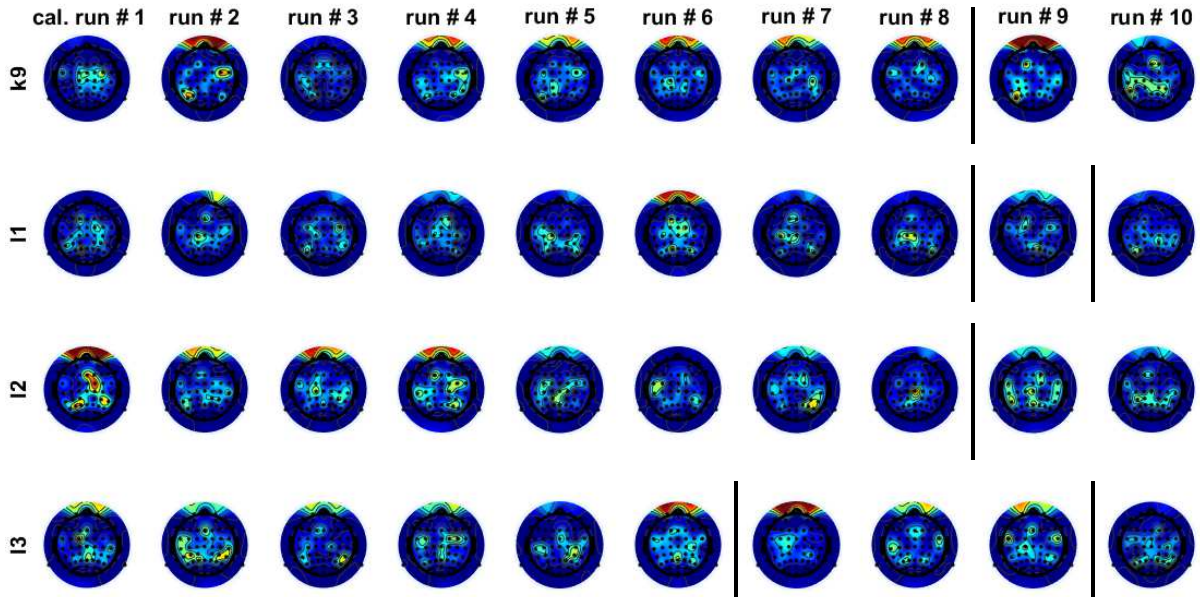


Figure 4.3: Fisher Scores computed on δ filtered calibration data, vertical lines correspond to a new session start

Figure 4.5, that depicts the scores computed on the β-band filtered signals reveals a different pattern: no participant appears to be able to produce completely stable features across the individual recordings, but in respect to the other two analyzed frequencies it yields higher significancy scores, especially in the case of participant K9 and L1. As the experiment evolves it appears that also for the β-band the hot spots tend to align toward the center midline. Notably, it is evident that the initial CPz FS observed in participants K9, L1 and L3 disappears in the runs belonging to the second session; consequently during the second session a second FCz peak seems to be registered for K9 and L3, possibly this arises from the introduction of the exoskeleton device and the physiological and mechanical alterations associated with the new constraints.

Evaluation The topographic maps corresponding to the evaluation runs capture distinct and subject-specific patterns for each frequency range considered.
Beginning with the lower section of the spectrum, the δ-band maps reported in Figure 4.6(a) reiterate the non-negligible role of centrally localized networks during motor imagery of a complete gait cycle. The plots of participants L2 and L3 show stable features at the frontal sites, with the lowest significance observed for participant K9, who displays a more posteriorly oriented focus. The near-complete disappearance of the significance index evaluated from Cz-originating signals is also noticeable.
The consistency observed in the calibration data, in terms of FS generated by motor oscillations μ, appears to be partially reduced in the evaluation data set. Most importantly, the maps (Fig-

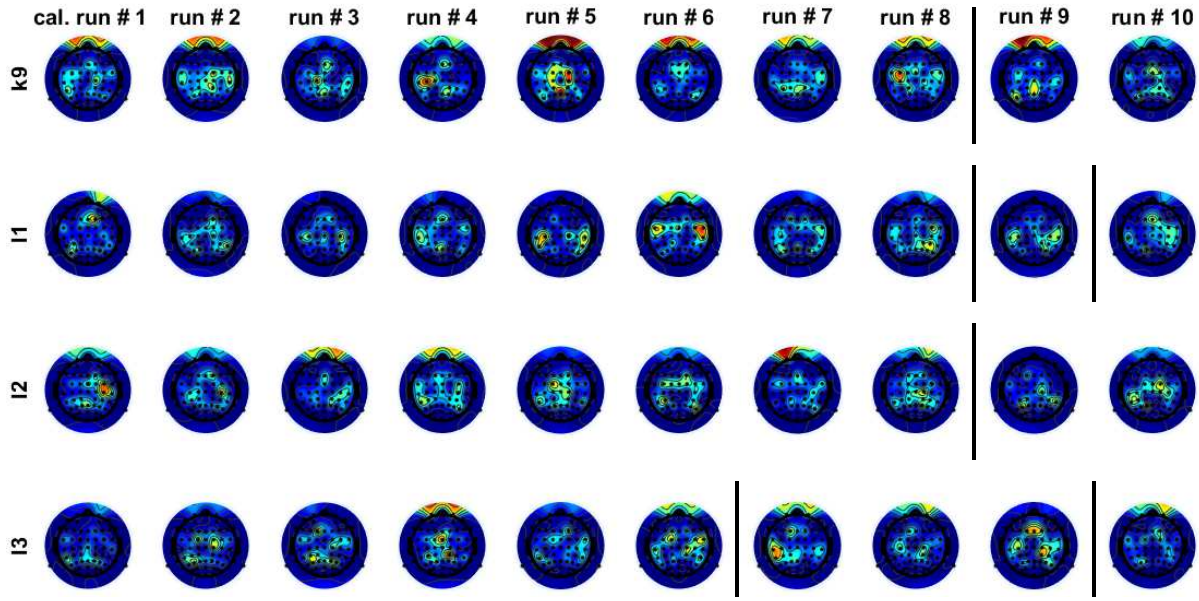


Figure 4.4: Fisher Scores computed on μ filtered calibration data, vertical lines correspond to a new session start

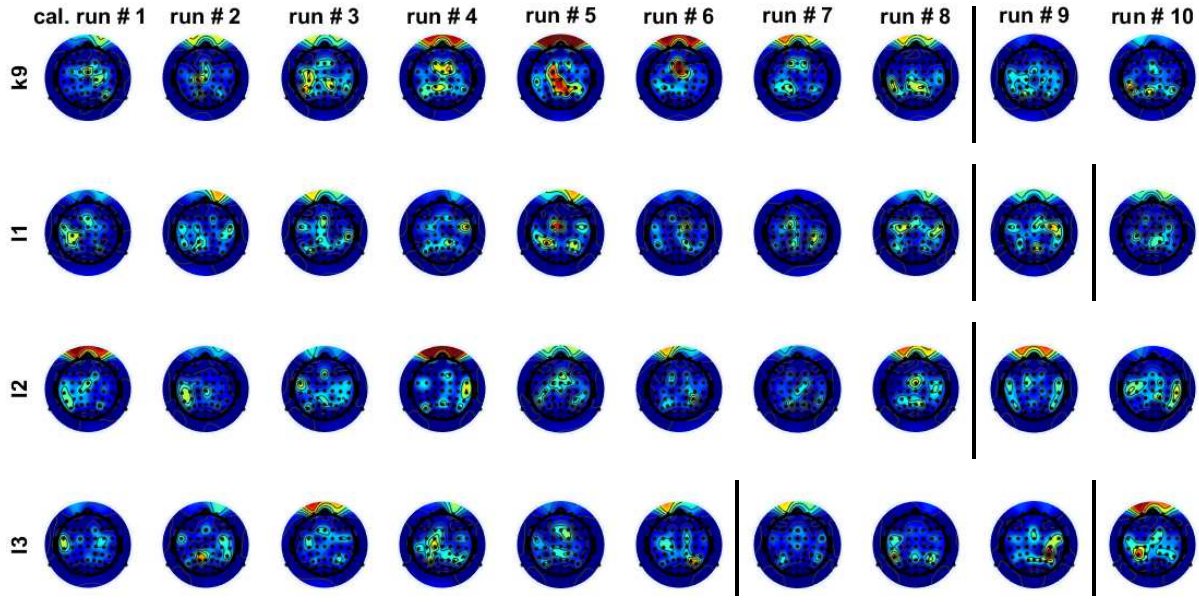
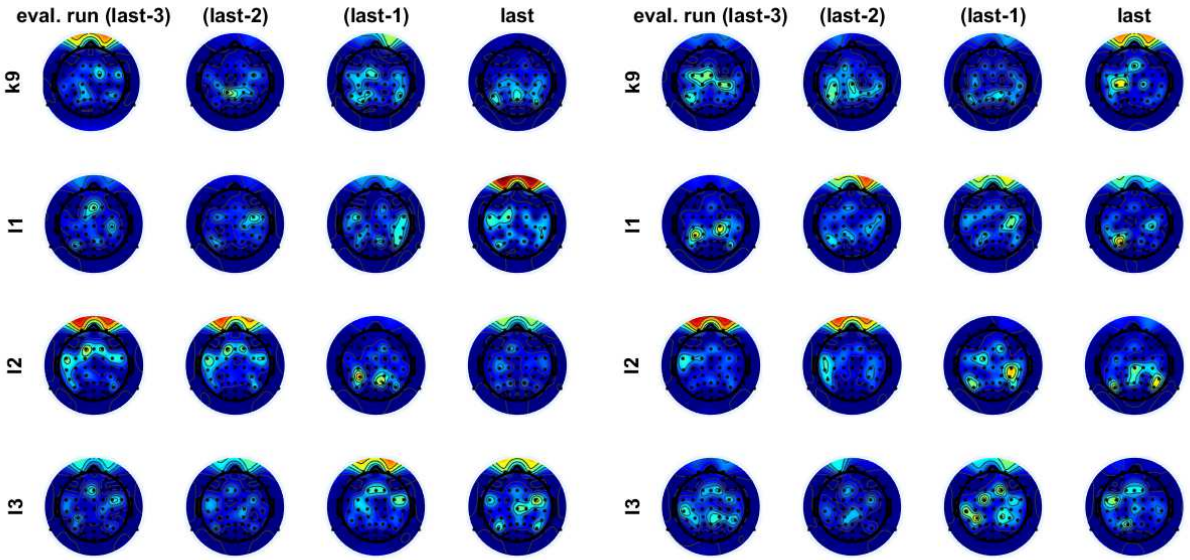


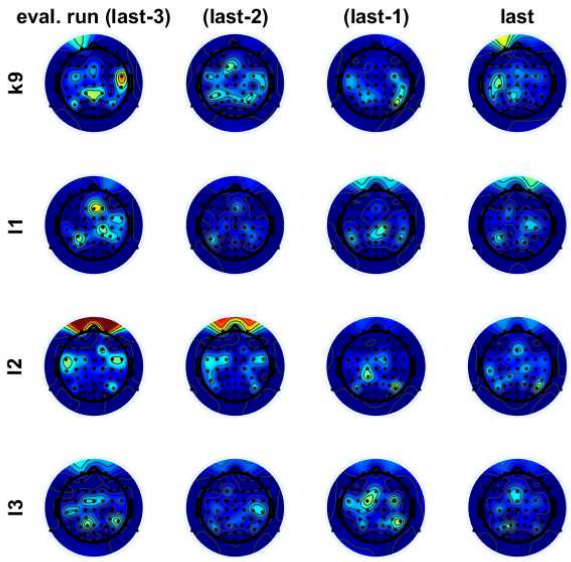
Figure 4.5: Fisher Scores computed on β filtered calibration data, vertical lines correspond to a new session start

ure 4.6(b)) indicate less frequent discriminant features arising from the frontal cortical region, which remains relevant mainly for subject L3.
Finally, the topoplots dispayed on Figure 4.6(c) computed from the filtered signals of the β band suggest that central posterior (for K9) and frontocentral (for L1 and L3) electrodes may have recorded relevant signals during evaluation runs. In the case of participant K9 and L1 this is consistent with the results of calibration data analysis.



(a)

(b)



(c)

Summary: FS maps calculated on the last 4 evaluation runs for the three frequency bands — δ (a), μ (b), and β (c).

Figure 4.6: FS maps of the last 4 evaluation runs

FS evolution A possible approach to assess the user's learning progression is to analyze the stability between runs of the score distributions [21]. In this framework, each recording file is represented as a 32-dimensional vector, whose elements correspond to the FS values computed across all EEG channels for that specific run. By quantifying the Euclidean distance between consecutive vectors, each representing temporally adjacent calibration or evaluation runs, it becomes possible to estimate the degree of temporal consistency or variability of the spatial FS pattern. A smaller distance indicates a stable and reproducible spatial distribution of discriminant information, suggesting consolidation of the user's neural activity, whereas larger distances may reflect continuous adaptation or signal variability.
Furthermore, to ensure cross-session and inter-subject comparability, each subject's FS vector was normalized with respect to the mean of the norms of all the participant-specific vectors. The resulting normalized distance values, reported in Figure 4.7, provide an interpretable measure of how the map of discriminative EEG features evolved throughout the experiment.

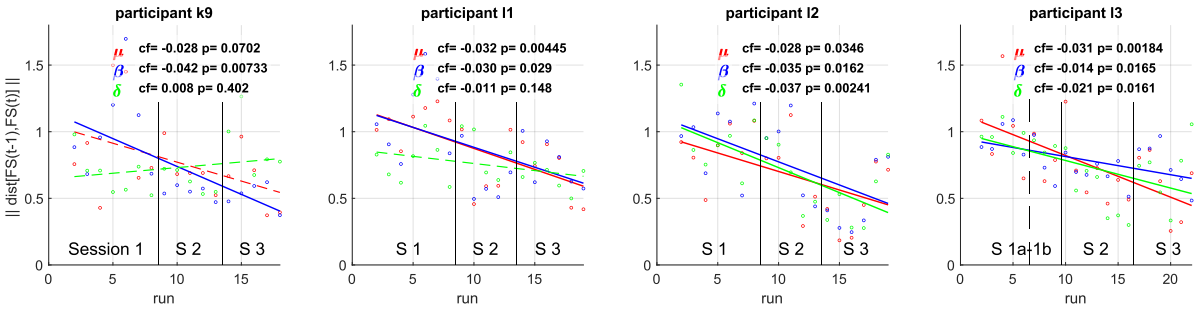


Figure 4.7: Distance between Fisher Scores vectors of two consecutive runs, normalized. For subject L3 the first session is divided into S1a and S1b because it took place in two different days.

The reported values denote a robust and widespread negative linear trend of run-to-run variability of the channels relevant for each frequency band. The common pattern shows that the variability between maps is high at the beginning of the protocol ($distance(v_t,v_{t-1}) > |\hat{v}|$) but, as the experiment evolves, the difference between the discriminative maps of two consecutive runs decreases, approaching values closer to one-half the mean FS-vector norm.
The only exceptions to this trend are recorded in the δ range for participant K9 and L1 and for participant K9 μ frequencies. In these cases, no positive trend is observed, but the negative correlation cannot be considered statistically significant (p-value > 0.05).
At the population level, given the available data, even before analyzing single-sample or single-trial outcomes, it can be assumed that internal user-learning processes appear to emerge consistently across subjects. The only frequency band interested for every participant in this process is the β band while the one showing less consistency is the δ band. A possible explanation is that, as reported by literature, the low-end of the spectrum is correlated to the gait exercise [49]

while the rest of the frequencies are closely linked to imagined motor activity and related learning processes [46], [47], [90]. In particular, the μ band reflects mechanisms in sensorimotor cortex related to motor imagery [47], while the β band is often considered an indicator of motor processing and attentional modulation during skill acquisition [90]. Since for the subjects the novelty lies in the mental exercise, and not in the performance of a complete gait cycle, it would therefore seem reasonable that the most evident improvements occur in the μ and β bands, consequently an increase in coherence within these bands can be interpreted as a sign of ongoing neural adaptation, consistent with existing literature on motor learning and cortical plasticity[90].

4.1.3 Single channel PSD

To provide further insight into the analysis of PSD evolution of recorded signals during the continuous feedback period, a time-frequency analysis was carried out on a restricted number of channels. Taking into account the previously reported results and the literature review, the electrodes selected to carry the time-frequency analysis are those lying on the midline, namely FCz, Cz, and Pz. To achieve a better visualization of the time-frequency spectra, a gaussian smoothing filter was performed.

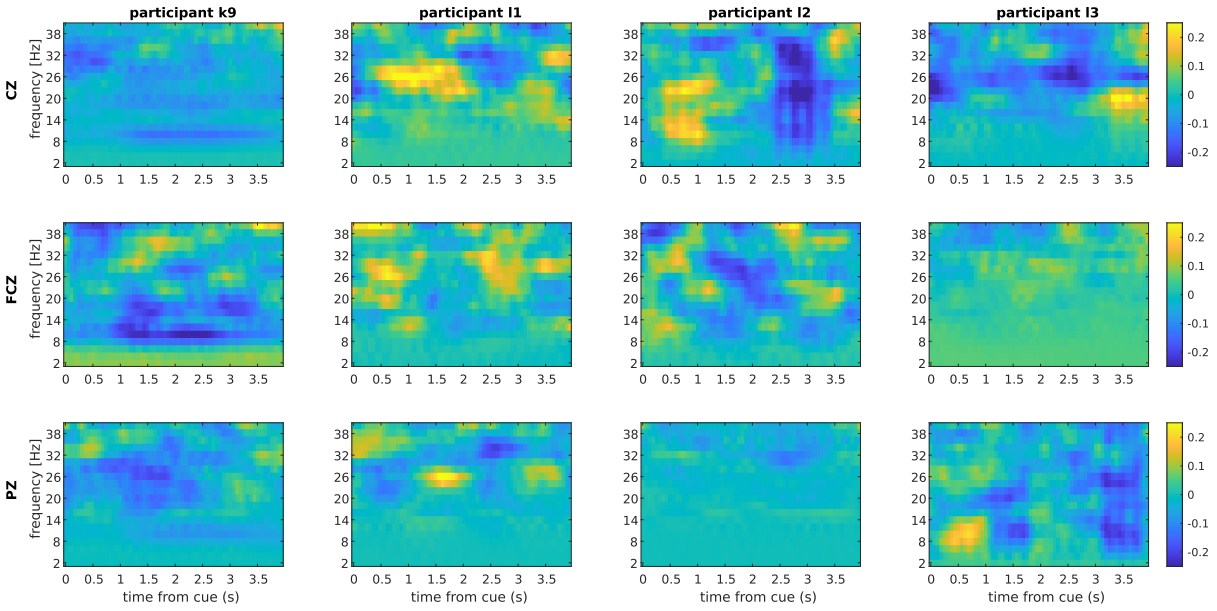


Figure 4.8: Time evolution of Power Spectral Density of movement-associated tasks referenced to the rest periods, calibration data

Calibration The evolution of the PSD is shown in Figure 4.8 (compared to the resting state, which is used as a reference period) estimated on calibration runs, reveals that the participant

producing the most sustained and notable desynchronizations (highlighted by blue colors) is K9, particularly on the frontal and central portion of the midline around the 10-12 Hz frequency range; posteriorly, as for the rest of the cohort, is the β frequency range the one subjected to similar alterations. The same long-lasting depression effect is detectable on the central electrode of participant L3 in the β interval.

The activity patterns depicted by participants L1 and L2, especially at the Cz and FCz site, are different and describe a time-dependent alternation of Synchronizations and Desynchronizations, which may resemble the movement-related potential fluctuations recorded during gait execution and reported by other studies [51].

For participant L1 and L2, the activity originating from the FCz location appears to follow the β rebound behavior (a delayed syncronization of the activity that occurs over the central electrodes after foot MI) described by Hashimoto and Ushiba [56].

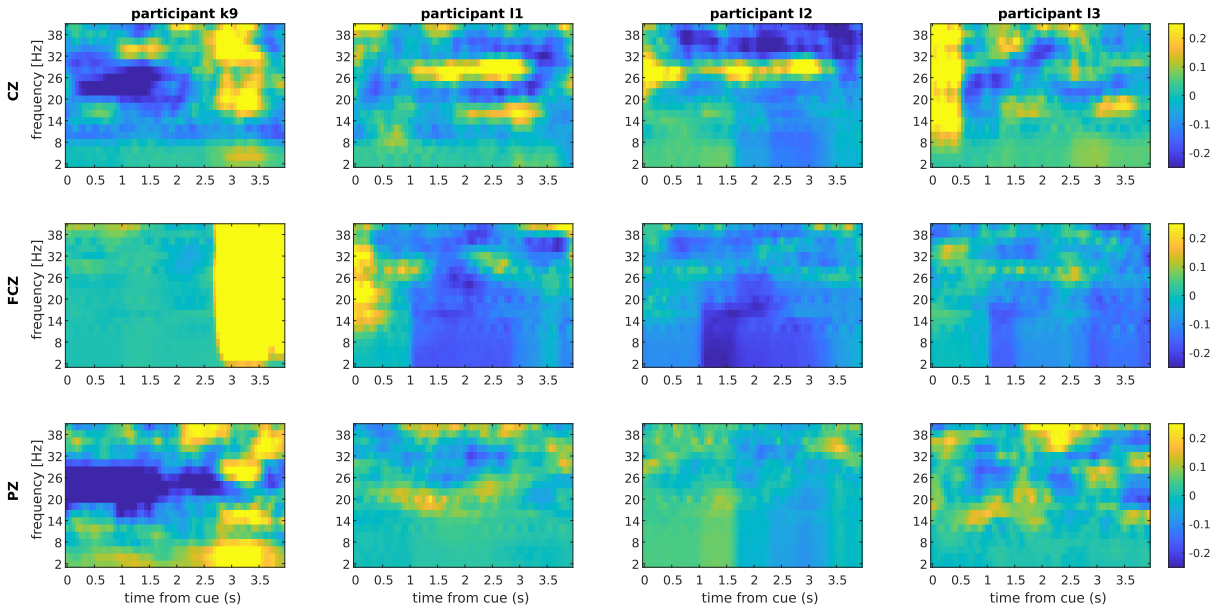


Figure 4.9: Time evolution of Power Spectral Density of movement-associated tasks referenced to the rest periods, evaluated on data from the last evaluation session

Evaluation The conduction of the same analysis on the runs acquired during the last session's online trials yielded coherent results reprorted by Figure 4.9. The most notable desynchronizations are recorded on participant's K9 posterior and central electrodes over a wide frequency band (20-30 Hz) that includes almost the whole β range.

In the δ and μ bands notable ERD can be observed over the fronto-central channel of the other participants.

For participant L2 it is also possible to observe the temporally constant high frequency central

desynchronization that exceeds the range of the frequency bands previously investigated and provides evidence for user-specific discriminant high frequency content.
Shifting the attention on population-level results, it is possible to infer that the channel-frequency feature pairs are highly user-specific and few common trend (outside the gradual spatial shift toward central sites) can be identified.

4.2 Classification models

For each participant, the adopted policy aimed to maintain the same unaltered classifier for as long as possible (unless unsatisfactory results were obtained). Accordingly, Figure 4.10 illustrates, for each subject and session, whether it was necessary to train a new classifier, perform a simple fine-tuning of the existing neural network with the new data, or whether the classifier from the previous session was preserved as it is.

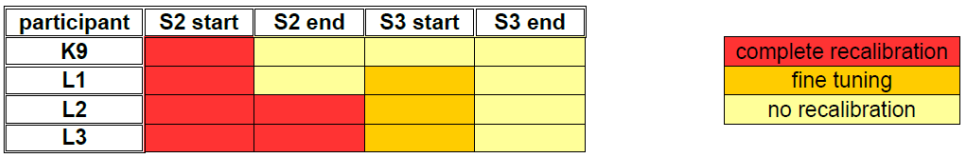


Figure 4.10: Classifier updates during evaluation sessions

Figure 4.11, on the other hand, shows the F1 scores achieved by all the classifiers used online when evaluated on the test set available at the time of training. In this context train, validation and test set are all independent prerecorded files divided approximately according to a 50% - 25% -25% ratio. It is important to note that these F1 scores were computed progressively, using only the data available up to each training point in the experiment. Therefore, although the test set is independent of the training and validation sets, the resulting scores may not fully reflect the online performance observed during the subsequent evaluation runs carried out with the same classifiers. Those performances are reported separately in the next subsection.
The main objective of the policy of infrequent calibrations is to promote the user-learning process. As can be inferred from both figures, this was achieved only for half of the participants (K9 and L1) whose initial classifiers proved to perform well on online applications. Differently, for participants L2 and L3, the neural networks required to be recalibrated at the beginning and end of the second session. Despite this, during the final session, no fine-tuning process was performed and all users were able to achieve positive results by adopting classifiers trained on the data sets gathered only in previous sessions.
The scores depicted in Figure 4.11 reflect the real-time evidence gathered during the experiment and reveal a marked user-specific variability in classifier performance.

Another noteworthy observation concerns the performance of PhiNet. Although never capable of achieving outstanding results, during the BCI refinement process, it proved to be the most reliable network for participants exhibiting lower overall performance. On the other hand, the remaining CNNs achieve peak performance for single subjects, but are not able to beat chance level classifiers for participant L1 and L2.

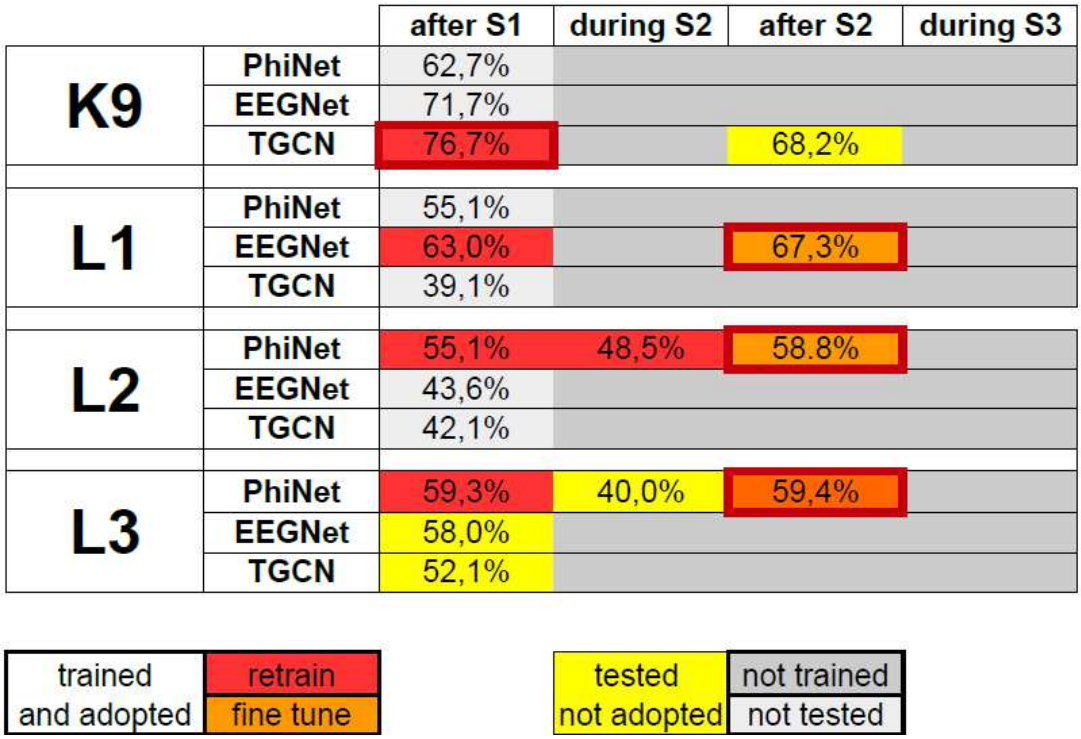


Figure 4.11: Performances of every classifier trained during (or before) evaluation sessions; classifiers lying in a red border cell are the one used for the last evaluation run of the third session

4.2.1 Models comparison

With the complete data set acquired during the entire duration of the experiment, it was possible to perform a posteriori offline data analysis regarding the performances of the single classifiers. As depicted in Figure 4.12, for each subject and each session, a fast training protocol (60 epoch, lr=0.0005 with linear decay to a 0.01 factor and patience of 15 epochs) was applied to compare the performance of every network with an ever increasing data set. For this experiment, the last two evaluation runs of the last session were utilized as a test set. As expected, a clear trend of improvements in the achieved results related to the widening of the data set can be observed. The best results arise from the predictions of Neural Network trained on the files recorded during the same acquisition session of the test set. Notable is the complete inability of the classifiers

		S1 runs	S1+S2 runs	S1+S2+S3 runs	mean
K9	PhiNet	53,7%	54,6%	54,6%	54,26% ± 0,53%
	EEGNet	56,6%	60,9%	66,5%	61,32% ± 4,98%
	TGCN	55,1%	68,3%	70,1%	64,47% ± 8,16%
L1	PhiNet	53,7%	43,9%	62,2%	53,26% ± 9,13%
	EEGNet	56,8%	59,9%	74,5%	63,74% ± 9,48%
	TGCN	44,1%	65,5%	68,8%	59,46% ± 13,38%
L2	PhiNet	29,8%	45,6%	59,1%	44,83% ± 14,69%
	EEGNet	38,4%	48,4%	56,8%	47,87% ± 9,19%
	TGCN	43,7%	45,6%	52,1%	47,10% ± 4,39%
L3	PhiNet	48,9%	43,3%	52,0%	48,09% ± 4,41%
	EEGNet	49,0%	53,2%	54,3%	52,16% ± 2,83%
	TGCN	47,0%	45,6%	45,5%	46,02% ± 0,87%
MEAN	PhiNet	46,50% ± 11,39%	46,86% ± 5,23%	56,97% ± 4,53%	50,11% ± 8,61%
	EEGNet	50,18% ± 8,63%	55,61% ± 5,91%	63,04% ± 9,29%	56,27% ± 9,15%
	TGCN	47,48% ± 5,29%	56,22% ± 12,33%	59,09% ± 12,25%	54,26% ± 10,80%

Figure 4.12: Performances (F1 score) of classifiers trained with data recorded during the three sessions and tested on the last two recorded files.

to identify reliably any pattern in the case of participants L2 and L3 when data of the same session are lacking. Simultaneously, the runs performed outside the exoskeleton (during the first calibration session) are generally not sufficient to train useful classifiers. The improving trend clearly highlights the session-to-session variability associated with this type of recording and reinforces the evidence of exoskeleton-mediated brain-patter modulations [52].

An interesting observation can be performed when these results are compared to the table of the previous subsection (Figure 4.11). Except for participant L3, the network architecture that produces the best results in the offline analysis coincides with the architecture adopted during the online runs. Of course, the observation is biased by the fact that the architecture employed online is selected as the one performing better on the dataset gathered and tested on the first session. Nevertheless, the fact that the participants are actively learning to adapt to the BCI-LLE system (while only one architecture is decoding their activity and providing the feedback) may polarize their modulation toward alterations detectable by the specific network.

From an architecture-selective perspective, the results obtained appear to align with the outcomes of the online classification analysis. When the data set is sufficiently representative to enable discriminative classification, PhiNet consistently achieves performances significantly higher than that of a random classifier (accuracy and F1-score > 51.8%, considering $p = 0.01$, $n = 4000$), although certain limitations still emerge as the network does not show a high session-to-session consistency. In contrast, TGCN demonstrates the capability to achieve good performances even in gait motor imagery tasks; however, its structural design does not allow it to generalize effectively across all participants.

Despite being the first architecture introduced, EEGNet proves to offer the most balanced trade-off between classification performance and generalization capability. This network maintains a robust F1-score across the subjects and achieves the best mean performance, suggesting its suitability as a reliable baseline model also for gait MI BCI studies. Nevertheless, the other two networks outperform EEGNet in certain cases, demonstrating that each architecture may capture subject-specific patterns that others overlook. Ignoring their contribution could therefore compromise the overall quality and adaptability of the system. For this reason, as previously suggested, it is not possible to identify a universal classifier that performs optimally for all subjects, highlighting the need for an individualized approach that can overcome the inter-subject variability [18].

4.2.2 Performance with different channels configuration

The final analysis conducted on the classifiers, using the complete dataset collected across the three recording sessions, aimed to evaluate the performance achievable when reducing the number of electrodes employed from the original 32-channel configuration. This comparison serves two main purposes. First, it allows verifying whether the trained neural networks actually base their pattern recognition processes on the electrodes previously identified as most informative during the offline analysis phase. Second, it enables testing the hypothesis that an excessive number of spatially and temporally correlated signals may introduce noise into the learning process. Under this assumption, a reduction in the number of input channels could encourage the architecture to focus on more salient and meaningful trends within the EEG data. To investigate these aspects, three spatial masks were applied iteratively to the EEG channels configuration, progressively excluding peripheral electrodes and reducing the number of input channels to 22, 16, and 10, all centered around the Cz electrode as shown in Figure 4.13 . Performance com-

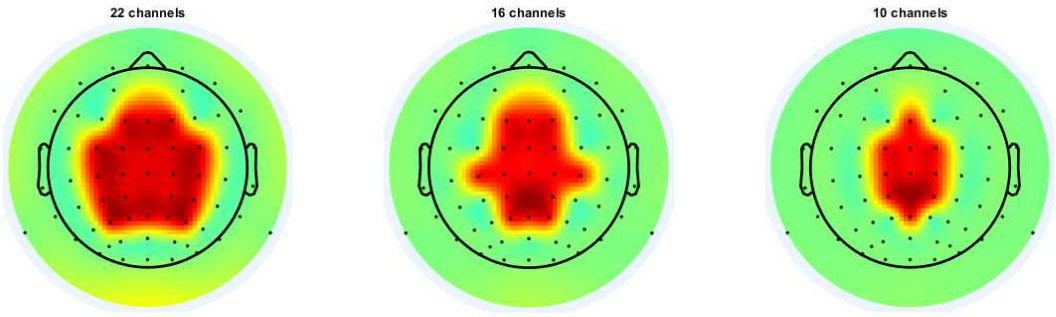


Figure 4.13: three masks utilized to selectively remove peripheral channels

parison across different electrode configurations highlights distinct behaviors among the three neural architectures considered. As reported by Figure 4.14, on average, EEGNet consistently

number of channels		32	22	16	10
K9	PhiNet	54,6%	60,5%	59,1%	51,3%
	EEGNet	66,5%	63,9%	62,5%	61,4%
	TGCN	70,1%	62,8%	61,4%	56,5%
L1	PhiNet	62,2%	59,7%	59,5%	52,7%
	EEGNet	74,5%	65,7%	66,4%	49,4%
	TGCN	68,9%	69,5%	74,7%	57,9%
L2	PhiNet	59,1%	47,3%	44,0%	50,1%
	EEGNet	56,8%	49,4%	54,9%	57,3%
	TGCN	52,1%	48,5%	50,9%	49,4%
L3	PhiNet	52,0%	44,6%	51,2%	54,7%
	EEGNet	54,3%	53,6%	53,3%	46,4%
	TGCN	45,5%	45,2%	48,2%	43,5%
MEAN	PhiNet	56,97% ± 4,53%	53,02% ± 8,21%	53,43% ± 7,35%	52,19% ± 1,96%
	EEGNet	63,04% ± 9,29%	58,12% ± 7,88%	59,27% ± 6,25%	53,63% ± 6,93%
	TGCN	59,12% ± 12,28%	56,49% ± 11,55%	58,80% ± 12,05%	51,80% ± 6,67%

Figure 4.14: Performances (F1 score) of classifiers trained with progressively reduced number of channels

achieved the highest accuracies in all configurations, with a mean accuracy ranging from 63.0% on 32 channels to 53.6% on 10 channels. This result suggests that EEGNet is relatively robust to channel reduction, maintaining a limited performance drop of about 9.4 percentage points even with a 70% decrease in spatial information.

TGCN followed a similar but slightly less stable trend, starting from 59.1% at 32 channels and declining to 51.8% with 10 channels. The smaller degradation indicates that the spatio-temporal graph convolutional mechanism can still extract meaningful features even when the spatial sampling is limited, possibly by taking advantage of the learned relationships between electrodes.

In contrast, PhiNet displayed more variable performance among participants, with accuracies oscillating around 57–52%. While the average reduction is moderate, the individual results (L2 and L3) reveal greater sensitivity to spatial information, suggesting that PhiNet's predictive mechanism benefits from denser input representations.

From a subject-wise perspective, L1 stands out as the most consistent performer, achieving top accuracies across all models and channel configurations, even reaching 74.7% with TGCN using only 16 electrodes, suggesting a notable capability of generalization and signal discriminability.

In contrast, L2 and L3 show overall lower and more unstable performances; yet it is interesting to underline the improvement of performance of EEGNet for participant L2 as the input size shrinks.

Overall, the results indicate that a moderate reduction in channel count (up to 16 electrodes centered on Cz) does not dramatically compromise decoding accuracy and, in certain cases (e.g.,

L1 with TGCN), may even lead to marginal improvements. This supports the hypothesis that excessive spatial redundancy may introduce noise rather than informative diversity, and that compact electrode configurations can enhance feature discriminability and model robustness. Furthermore the results are consistent with the feature analysis previously performed.

4.3 Online evaluation results

Figure 4.15 summarizes the results obtained during the first and last runs of the two evaluation sessions.

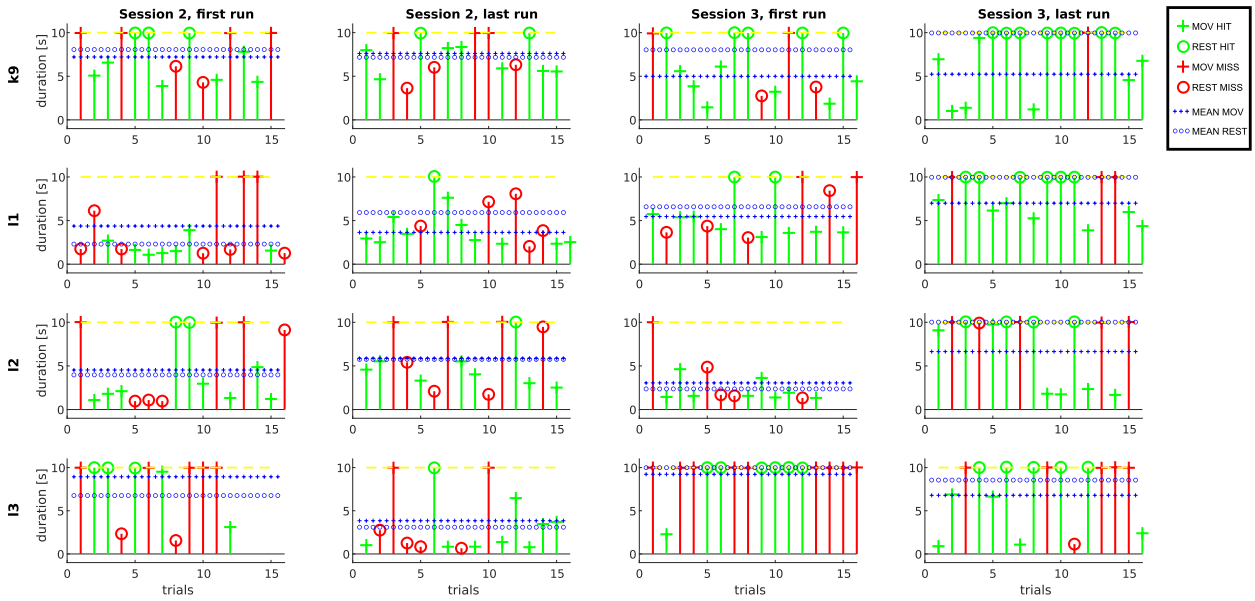


Figure 4.15: Single trial outcomes of online evaluation runs. For each evaluation session the first and the last run are considered. Green labeled stems represent a positive outcome while red stem correspond to failed attempts.

The first clearly observable trend is the progressive left-to-right gradient toward increasingly green-saturated maps, which underlines a general improvement in single-trial outcome accuracy. The second noteworthy aspect emerges from the analysis of the mean times spent in movement- and rest-related continuous feedback epochs, represented by blue dotted horizontal lines (o = Rest; $+$ = Movement). An effective BCI should promote rapid recognition of movement intention and a stable and sustained management of rest states; therefore, the mean rest time should exceed the mean MI time, and the greater this difference, the better the system's performance. As shown, all participants, except for K9, initially exhibited negative results during the first evaluation session, which gradually improved as the experiment evolved. The clear separation between the two mean values observed during the last run of the final session further

supports this trend. This outcome suggests that a three-session training protocol may already be sufficient to achieve meaningful improvements in user performance.
From the same graphs, it is also interesting to note how the performance outcome generally improved during each evaluation session and also how it changed significantly between the end of the second and the start of the third session.
An important consideration that emerges by analyzing Figure 4.15 and considering the feedback of the participants is the fact that the nature of the gait movement imagination exercise does not favor a continuous type of control. In particular, participants L2 and L3 reported a marked difficulty in controlling the BCI system if it failed to recognize the first Step cycle imagination. As the last column of plots reports for both of them, the movement command threshold was either quickly reached or not crossed at all, and only in few instances a delayed command was possible (trial 1 for L2; trials 2 and 4 for L3).
Table 4.1 reports the values of the hyperparameter ev (that controls the speed of the feedback bar and th (that acts as an evidence threshold) used during each evaluation run of the second and third sessions. The consideration of these values is helpful for the interpretation of the analytical metrics summarized in Figure 4.16 that describe trial outcomes.

	K9		L1		L2		L3	
	ev	th	ev	th	ev	th	ev	th
run 1 S2	56	0.8	48	0.7	48	0.7	56	0.7
run 2 S2	48	0.8	48	0.8	48	0.7	48	0.6
run 3 S2	48	0.8	48	0.8	56	0.8	48	0.7
run 4 S2	48	0.8	48	0.8	–	–	48	0.65
run 5 S2	–	–	–	–	–	–	48	0.65
run 6 S2	–	–	–	–	–	–	48	0.65
run 1 S3	48	0.8	56	0.8	48	0.75	48	0.8
run 2 S3	48	0.8	56	0.8	48	0.8	48	0.7
run 3 S3	48	0.8	48	0.8	48	0.85	48	0.75
run 4 S3	48	0.8	48	0.8	56	0.8	56	0.8
run 5 S3	–	–	48	0.8	56	0.85	56	0.7
run 6 S3	–	–	–	–	56	0.8	–	–

Table 4.1: Hyperparameter values used during online evaluation runs

Across the population, the evaluation metrics reveal heterogeneous, yet interpretable, learning trends throughout the experimental sessions. When data is aggregated over all subjects, a general tendency toward improved trial classification accuracy can be observed over time, although the extent and statistical reliability of these improvements differ considerably between individuals.
For most of the participants (K9, L1, and L3) both trial-level and sample-level accuracies exhibit

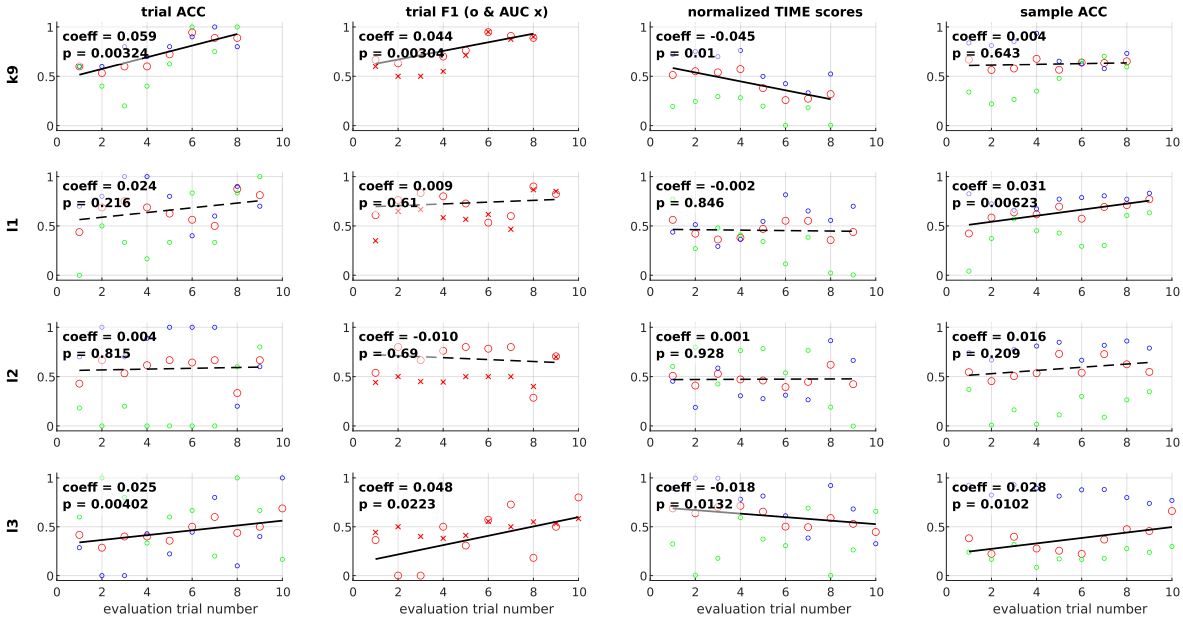


Figure 4.16: single trial results for all the evaluation runs performed, red dotted marks depict the values reported by the title and fitted by the regression line, blue dots are relative to movement-class metrics, green dots to rest-class metrics

positive improvements, sometimes statistically significant ($p < 0.05$). This pattern suggests that, under repeated exposure, users progressively refined their ability to produce more distinguishable neural patterns. Simultaneously the classifiers were able to decode with increasing stability the brain activity patterns. In contrast, subject L2 exhibited no significant trend in any of the accuracy-related metrics, suggesting limited adaptation with training.

Of interest is the fact that the normalized time score, which reflects task efficiency (with lower values representing faster or more stable control), tends to decrease across sessions for most of the participants (no increase is registered) and shows the highest significance in the case of participants who underwent accuracy improvements. The simultaneous improvement in both sample accuracy and timing performance suggests a genuine learning effect rather than random variability.

From an individual perspective, K9 stands out as the most consistent learner, showing significant positive trends in trial metrics (both accuracy and F1 scores regression shows p-value < 0.05), accompanied by a clear improvement of the normalized time score. The most outstanding consideration regarding these outcomes is the fact that during all the experiments no recalibration process and only one hyperparameter change was performed. Despite being counterintuitive, it is possible for trial-based accuracy to improve even if single-sample accuracy does not. This occurs because the evidence integration process introduces temporal correlation between classifier samples; if the trial features become more stable, erroneous classifications tend to concentrate

in a few misclassified trials rather than being spread across all trials. Furthermore, if the subject learns to produce more distinguishable modulations, the average classifier confidence improves and the integrator moves faster toward the correct class. Therefore, it is possible to attribute the improvements to the learning process of the participant.

L1 presents a similar, though milder, behavior, with significant improvements in sample accuracy and only few hyperparameter adjustments. The user's ability to adapt and learn led, in the last two runs, to outstanding results: driven by the highest recorded sample accuracy of the cohort, on these two repetitions the participant reached trial-based F1 scores of 90.0% and 82.4% characterized by a near perfect rest task accuracy.

L2 remains substantially flat across all measures, reinforcing the notion of pronounced inter-subject variability, a well-documented feature of MI-based BCI training [17], [21]. However, unlike other participants, L2 started from relatively high baseline performance, already exceeding the chance level expected from a random classifier. This outcome was driven by the fact that the neural patterns corresponding to the movement imagery class were well discriminated from the beginning, as reported by the blue dots of the single sample accuracy graph of Figure 4.16. Yet, this apparent advantage introduced a critical limitation: the online trial-based performances of L2 showed a strong movement-class polarization, making it difficult for the participant to sustain the "rest" state during control tasks. As a result the user struggled to complete several control trials and, to mitigate this situation, specific attention was directed toward tuning the integrator hyperparameters. By gradually increasing the decision threshold and regulating the speed parameter, more balanced results were achieved. Notably, during the last two runs, the rest-class metrics finally departed from the baseline values showing a marked improvement that enabled the participant to exceed for the first time the AUC value expected from a random classifier. With a larger number of sessions, it is plausible that this participant would have shown a positive performance trend comparable to that of the others.

Participant L3 also represents a peculiar case study: despite not being able to achieve chance-level classifiers metrics in the beginning, thanks to the steady and relevant increase of classifier accuracy, the user was able to achieve satisfactory advancement in controlling the feedback bar reflected by the steady accuracy and time-score improvement.

Overall, these results highlight a population-level progression toward higher classification performance and reduced response times in a subset of users. The coexistence of improving accuracy and decreasing normalized time scores in most participants is consistent with the development of a more stable and discriminative neural representation of motor imagery, supporting the hypothesis that user learning is a critical component to consider to the development of an LLE-MI-BCI system.

Chapter 5

Conclusions

The experimental results obtained confirm that the proposed protocol is capable of supporting the development of a functional BCI system while simultaneously promoting neural learning processes in the user. The combined analysis of Fisher Score maps, ERD/ERS patterns, and spectral power distributions reveals a progressive centralization of discriminant information within the cortical regions over the course of the sessions. This phenomenon, together with the shift of power peaks in midline channels toward physiologically meaningful frequency bands (μ and β), suggests the emergence of progressive cortical specialization associated with the user's familiarization with the motor imagery task. In parallel, the widespread improvement in reaction times and the increase in online test accuracy metrics further support the hypothesis that the protocol promotes genuine BCI learning, in which user and system adapt to each other reciprocally.
Another noteworthy aspect concerns the selection of the classifier. The experience gained throughout the experiment indicates that no single network or architecture is universally superior; rather, the choice of model constitutes an actual hyperparameter of the system, whose effectiveness is closely dependent on the individual characteristics of the signal. EEGNet demonstrated the best inter-subject generalization capability, whereas TGCN and PhiNet exhibited higher performance in specific cases, showing that different approaches can capture complementary patterns. This observation supports the perspective of adaptive BCI systems, in which the classifier can be personalized or co-evolve with the user.
The reduction of the number of EEG channels yielded equally significant results. While confirming that 32 electrodes ensure maximal spatial coverage, the comparative analysis showed that more compact configurations (down to 16 channels centered on Cz) maintained comparable performance, in some cases even improving model stability and robustness. This finding suggests that an excess of spatially correlated signals may, in isolated instances, introduce noise or redundancy, thus hindering the machine learning process. Nevertheless, the possibility of drastically reducing the number of channels must be approached with caution: Fisher Score maps

and ERD/ERS variations indicate that cortical patterns are not static but evolve substantially from session to session, particularly in the presence or absence of the exoskeleton. Classifiers trained on previous sessions, especially for subjects L2 and L3, did not always generalize well as session passes, highlighting the strong dependence between cognitive state, and EEG patterns.
Overall, the collected evidence shows that the application of the proposed approach, based on repeated sessions, continuous feedback, and limited calibration, can serve as an effective strategy to stimulate user learning and consolidate brain control in BCI contexts for motor substitution or rehabilitation assisted by an exoskeleton. At the same time, the results highlight the need for a dynamic evolution of the classification methodologies capable of adapting not only to individual peculiarities, but also to the temporal evolution of the user's neural networks.
In future work, longer-term studies with a larger number of participants may further clarify the relationship between learning dynamics, cortical variations, and decoder performance. Nevertheless, the results obtained demonstrate that the proposed pathway is promising: a BCI-LLE system not only can become progressively more accurate in recognizing motor intentions, but it can also enhance the neural learning of the user, transforming the control process into an adaptive bidirectional dialog between brain and machine.
With appropriate attention to the training process and to the implementation of a shared control system between the robot and the participant, it is conceivable that such an approach could enable the control of a lower-limb exoskeleton in continuous operation modes, not constrained by predetermined cue sequences, but entirely driven by the user's motor intention; to allow this type of control, however, it is necessary to achieve an almost flawless classification of the rest state, also in more complex environments.

5.1 Limitations and future developments

The conducted study, although capable of achieving positive results, is affected by several limitations.
Firstly, the size of the participant cohort is relatively small and clearly not large enough to allow population-level studies to investigate possible correlations between user-specific parameters and the outcomes obtained. For instance, it would be interesting to explore the hypothesis that participants with prior football experience (such as K9) may be more likely to produce more distinct foot-related activity modulations.
Secondly, the three-session structure adopted for the experiment might have limited the ability to identify learning patterns in certain participants. In particular, participant L2 (who, despite acceptable classifier performance, required an extensive parameter optimization process to achieve good results in the final runs) might have shown clearer improvement trends if one or more ad-

ditional sessions had been included. For the rest of the cohort, additional sessions could also have allowed an assessment of the stability of the results obtained.

A last consideration regarding the cohort is the fact that the system has been tested on healthy volunteers and needs to be validated by motor impaired end users. As reported by Tonin & Millán [18] this may raise an important translation gap and would likely require additional investigations.

Furthermore, it should be emphasized that, in order to reduce the variability of online results and to facilitate a faster approach to the evaluation sessions, it was necessary to minimize the number of selected classifiers and to consider only a single basic probability integrator. More complex algorithms, especially those based on dynamical systems[89], could potentially yield better performance in terms of single-trial accuracy, but require a more accurate and time-consuming process of hyperparameter selection.

Finally, it should be addressed that the proposed framework enables the possibility to acquire right-step and left-step labeled data to allow potential post-hoc analyses or future reuse of the setup for more detailed lateralization studies. During our experiment no distinction was made during the evaluation phase as the classifiers were trained to decode the general 'Step' motor command and the correct motor instruction was generated by a simple logical circuit. As discussed before, the distinction between right and left foot motor imagery is complex and would have yielded unsatisfactory results for beginner participants. Although this may appear as a substantial simplification, it remains a meaningful approach when considering the sequential nature of gait. This structure also enabled the development of a straightforward step algorithm for the exoskeleton, demonstrating that even with a simplified design a functional interface could be achieved. Nevertheless, it would be interesting to analyze the possibility to discern left and right step intention after a larger number of sessions if the participants establish a constant performance with the simple Rest-Movement exercise.

Bibliography

[1] S. A. et al., "Prevalence and incidence of ictus in europe: Systematic review and meta-analysis," 2022.

[2] W. H. et al., "Burden of stroke in europe: Thirty-year projections of incidence, prevalence, deaths, and disability-adjusted life years.," 2020.

[3] E. H. et al., "A review of rehabilitative and assistive technologies for upper-body exoskeletal devices," 2023.

[4] L. E. M. et al., "Clinical effectiveness and safety of powered exoskeleton-assisted walking in patients with spinal cord injury: Systematic review with meta-analysis," 2016.

[5] E. P. et al., "Mobility impact and well-being in later life: A multidisciplinary systematic review," *Research in Transportation Economics*, 2021.

[6] L. M. et al., "The-state-of-the-art of soft robotics to assist mobility: A review of physiotherapist and patient identified limitations of current lower-limb exoskeletons and the potential soft-robotic solutions," 2023.

[7] C. T. et al., "A review of soft wearable robots that provide active assistance: Trends, common actuation methods, fabrication, and applications," *Cambridge University Press*, 2020.

[8] C. D. et al., "Efficacy of robotic exoskeleton for gait rehabilitation in patients with subacute stroke : A systematic review.," 2022.

[9] Y. J. et al., "Effect of exoskeleton robot-assisted training on gait function in chronic stroke survivors: A systematic review of randomised controlled trials.," 2023.

[10] A. J. et al., "Stride management assist exoskeleton vs functional gait training in stroke a randomized trial.,"

[11] M. G. et al., "Lower limb exoskeleton robot and its cooperative control: A review, trends, and challenges for future research.," 2023.

[12] C. M. et al., "Brain-computer interfaces for post-stroke motor rehabilitation: A meta-analysis.," 2018.

[13] X. L. et al., “Effects of motor imagery based brain-computer interface on upper limb function and attention in stroke patients with hemiplegia: A randomized controlled trial,” 2023.

[14] Z. M. et al., “Motor imagery-based brain–computer interface rehabilitation programs enhance upper extremity performance and cortical activation in stroke patients,” 2024.

[15] F. A. et al., “Post-stroke rehabilitation training with a motor-imagery-based brain-computer interface (bci)-controlled hand exoskeleton: A randomized controlled multicenter trial.,” 2017.

[16] B. N. et al., “Neural activity modulations and motor recovery following brain-exoskeleton interface mediated stroke rehabilitation.,” 2020.

[17] P. et al., “Brain-machine interfaces: A tale of two learners,” *IEEE Systems, Man, and Cybernetics Magazine*, 2020. doi: 10.1109/MSMC.2020.2992776.

[18] T. Millán, “Noninvasive brain–machine interfaces for robotic devices. annual review of control, robotics, and autonomous systems,” 2021.

[19] K. et al., “The first communications about operant conditioning of the eeg,” 2011.

[20] P. et al., “The cybathlon bci race: Successful longitudinal mutual learning with two tetraplegic users,” 2018.

[21] T. et al., “Neural correlates of user learning during long-term bci training for the cybathlon competition.,” 2022.

[22] S. W. et al., “Invasive vs. non-invasive neuronal signals for brain-machine interfaces,” 2016.

[23] S. G. et al., “Recoding two-dimensional movement trajectories using electrocorticographic signals in humans.,” 2007.

[24] L. et al., “The past present and future of in vivo-implantable recording microelectrodes: The neural interfaces,” 2024.

[25] P. V. S. et al., “Response of brain tissue to chronically implanted neural electrodes,” 2005.

[26] A. N. G. et al., “Biomaterial strategies for regulating the neuroinflammatory response,” 2023.

[27] S. W. et al., “A review on directional information in neural signals for brain-machine interfaces,” 2009.

[28] Y. D. et al., “Neural decoding for intracortical brain–computer interfaces,” 2023.

[29] C. I. et al., “Brain-computer interface control of stepping from invasive electrocorticography upper-limb motor imagery in a patient with quadriplegia,” 2023.

[30] L. J. et al., “Real-time brain-computer interface control of walking exoskeleton with bilateral sensory feedback,” 2025.

[31] F. Emanuela, *Lecture notes for the course ”bioengineering for neurorehabilitation 2024-2025”*, Teaching material, University of Padova, Prof. Formaggio Emanuela, 2025.

[32] abian Pedregosa-Izquierdo, “Feature extraction and supervised learning on fmri: From practice to theory,” 2015.

[33] W. kin Tam et al., “Human motor decoding from neural signals: A review,” 2019.

[34] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience: Exploring the Brain*. Baltimore: Lippincott, 2001.

[35] H. B. et al., “Optimization of ssvep brain responses with application to eight-command brain–computer interface.,” 2010.

[36] S. S. et al., “Evoked-potential correlates of stimulus uncertainty.,” 1965.

[37] G. M. et al., “A metric for thought: A comparison of p300 latency and reaction time,” 1981.

[38] F. L. A. et al., “Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials.,” 1988.

[39] *Brainaccess website*, `https://www.brainaccess.ai/a-newly-developed-fast-p300-classifier/`, Accessed: 2025-11-10.

[40] F. P. et al., “Error-related eeg potentials generated during simulated brain-computer interaction.,” 2008.

[41] R. C. et al., “Robust, accurate spelling based on error-related potentials,” 2016.

[42] E. M. W. et al., “Noninvasive mapping of muscle representations in human motor corte,” 1992.

[43] F. J. et al., “Accounting for stimulations that do not elicit motor-evoked potentials when mapping cortical representations of multiple muscles,” 2022.

[44] W. PENFIELD and E. BOLDREY, “Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation,” 1937.

[45] W. H. et al., “Occipital alpha-band brain waves when the eyes are closed are shaped by ongoing visual processes,” 2022.

[46] J. W. C. et al., “Beta-band activity and connectivity in sensorimotor and parietal cortex are important for accurate motor performance,” 2016.

[47] R. H. et al., “Magnetoencephalographic cortical rhythms,” 1997.

[48] S. A. et al., ""eeg-based bcis on motor imagery paradigm using wearable technologies: A systematic review," 2023.

[49] T. et al., "Hybrid human-machine interface for gait decoding through bayesian fusion of eeg and emg classifiers.," 2020.

[50] Y. et al., "Neural decoding of gait phases during motor imagery and improvement of the decoding accuracy by concurrent action observation.," 2021.

[51] T. et al., "Deep learning-based bci for gait decoding from eeg with lstm recurrent neural network," 2020.

[52] T. et al., "Effect of lower limb exoskeleton on the modulation of neural activity and gait classification," 2023.

[53] H. Y. et al., "Negative covariation between task-related responses in alpha/beta-band activity and bold in human sensorimotor cortex: An eeg and fmri study of motor imagery and movements," 2011.

[54] T. et al., "Mu-beta event-related (de)synchronization and eeg classification of left-right foot dorsiflexion kinaesthetic motor imagery for bci.," 2020.

[55] W. et al., "Somatotopic arrangement of the human primary cortex," 2021.

[56] J. U. Yasunari Hashimoto, "Eeg-based classification of imaginary left and right foot movements using beta rebound," 2013.

[57] L. F. N.-A. et al., "Brain computer interfaces, a review," 2012.

[58] W. S. M. W. Pitts, "A logical calculus of the ideas immanent in nervous activity.," *ulletin of mathematical biophysics*, 1943.

[59] M. I. et al., "An overview of neural network," 2019.

[60] Z. C. L. et al., "A critical review of recurrent neural networks for sequence learning," 2015.

[61] V. J. L. et al., "Eegnet: A compact convolutional neural network for eeg-based brain-computer interfaces," 2018.

[62] Z. L. et al., "A survey of convolutional neural networks: Analysis, applications, and prospects," 2022.

[63] B. D. et al., "Deep learning approach for eeg classification in lower-limb movement phases: Towards enhanced brain-computer interface control.," 2024.

[64] S. N. et al., "An empirical comparison of neural networks and machine learning algorithms for eeg gait decoding.," 2020.

[65] R. T. S. et al., “Deep learning with convolutional neural networks for eeg decoding and visualization,” 2017.

[66] J. L. Elman, “Finding structure in time,” 1990.

[67] N.-S. K. et al., “A lower limb exoskeleton control system based on steady state visual evoked potentials.,” 2015.

[68] F. et al., “Brain symmetry analysis during the use of a bci based on motor imagery for the control of a lower-limb exoskeleton.,” 2021.

[69] L. et al., “A brain-controlled exoskeleton with cascaded event-related desynchronization classifiers.,” 2017.

[70] K. et al., “Eeg differentiates left and right imagined lower limb movement.,” 2021.

[71] M. et al., “Eeg signals classification and determination of optimal feature-classifier combination for predicting the movement intent of lower limb,” 2016.

[72] B. et al., “Performance analysis of left and right lower limb movement classification from eeg.,” 2016.

[73] C. et al., “Developing a motor imagery-based real-time asynchronous hybrid bci controller for a lower-limb exoskeleton,” 2020.

[74] W. et al., “A control system of lower limb exoskeleton robots based on motor imagery,” 2017.

[75] L. et al., “The human–machine interface design based on semg and motor imagery eeg for lower limb exoskeleton assistance system.,” 2024.

[76] T. et al., “Classification of left and right knee extension motor imagery using common spatial pattern for bci applications.,” 2019.

[77] *Indi global website*, `https://www.indi.global/home`, Accessed: 2025-11-10.

[78] *Ant neuro website*, `https://www.ant-neuro.com/`, Accessed: 2025-11-20.

[79] K. Anis, *Robot Operating System (ROS)*. Springer, 2017.

[80] *Ros-neuro github repository*, `https://github.com/rosneuro`, Accessed: 2025-11-12.

[81] C. et al., “The effect of user learning for online eeg decoding of upper-limb movement intention,” 2025.

[82] A. et al., “Effect of biased feedback on motor imagery learning in bci-teleoperation system,” 2014.

[83] B. S, “On the theory of filter amplifiers,” 1930.

[84] M. Federica, “Sviluppo di una self-paced bmi per l’identificazione dell’intenzione di movimento ed il controllo di un esoscheletro per il cammino,” M.S. thesis, Università degli studi di Padova, 2025.

[85] P. et al., “Event-related eeg/meg synchronization and desynchronization: Basic principles,” *Clinical Neurophysiology*, 1999.

[86] I. C. C. et al., “Temporal graph convolutional networks for automatic seizure detection,” 2019.

[87] S. I. et al., “Phinets: Brain-inspired non-contrastive learning based on temporal prediction hypothesis,” 2025.

[88] Y. C. et al., “Predictive sequence learning in the hippocampal formation,” 2024.

[89] T. L. et al., “The role of the control framework for continuous teleoperation of a brain–machine interface-driven mobile robot.,” 2019.

[90] S. E. et al., “Sensorimotor cortex beta oscillations reflect motor skill learning ability after stroke,” 2020.