Motor imagery classification in a brain-computer interface using kinesthetic vibrotactile guidance and information entropy

Batistić, Luka

UNIVERSITY OF RIJEKA
FACULTY OF ENGINEERING

Luka Batistić

MOTOR IMAGERY CLASSIFICATION IN A BRAIN-COMPUTER INTERFACE USING KINESTHETIC VIBROTACTILE GUIDANCE AND INFORMATION ENTROPY

DOCTORAL DISSERTATION

Rijeka, 2023.

UNIVERSITY OF RIJEKA
FACULTY OF ENGINEERING

Luka Batistić

MOTOR IMAGERY CLASSIFICATION IN A BRAIN-COMPUTER INTERFACE USING KINESTHETIC VIBROTACTILE GUIDANCE AND INFORMATION ENTROPY

DOCTORAL DISSERTATION

Supervisor: Assoc. Prof. Sandi Ljubić, PhD
Co-supervisor: Prof. Ivan Štajduhar, PhD

Rijeka, 2023.

SVEUČILIŠTE U RIJECI
TEHNIČKI FAKULTET

Luka Batistić

KLASIFIKACIJA MOTORIČKIH PREDODŽBI U SUČELJU MOZAK-RAČUNALO POMOĆU KINESTETIČKOGA VIBROTAKTILNOG NAVOĐENJA I INFORMACIJSKE ENTROPIJE

DOKTORSKI RAD

Rijeka, 2023.

Supervisor: Assoc. Prof. Sandi Ljubić, PhD (University of Rijeka, Faculty of Engineering)
Co-supervisor: Prof. Ivan Štajduhar, PhD (University of Rijeka, Faculty of Engineering)

Doctoral dissertation was defended on _____ at University of Rijeka, Faculty of Engineering, in front of the committee:

1. Prof. Ivo Ipšić, PhD - committee chair (University of Rijeka, Faculty of Engineering)
2. Prof. Miroslav Vrankić, PhD - member (University of Rijeka, Faculty of Engineering)
3. Assist. Prof. Peter Rogelj, PhD - member (University of Primorska - Faculty of Mathematics, Natural Sciences and Information Technologies, Slovenia)

ACKNOWLEDGEMENTS

Sincere thank you to my mentor, Prof. Sandi Ljubić, my co-mentor Prof. Ivan Štajduhar, and my friend and colleague Diego Sušanj, who were there for me when I was in dire straits.

Many thanks to all my colleagues who have accompanied me on this journey and provided wise counsel. My special thanks go to the team of the Institute of Neural Engineering at Graz University of Technology. It was a great honor to work with Prof. Gernot R. Müller-Putz, Prof. Andreea Ioana Sburlea, and Lea Hehenberger.

Last but not least, I would like to thank my friends and family for brightening up my days and making the time fly by. Nikola, Mimi, Mom, Dad and Mario, thank you for all the support and love. Ana and Roza, you are my everything and you made this possible.

ABSTRACT

Motor imagery (MI), a cognitive process involving the imagined performance of a motor task without actual movement execution, plays an important role in assistive technologies, rehabilitation, and the control of brain-computer interfaces (BCIs) that use electroencephalographic (EEG) recordings. This dissertation emphasizes the application of innovative approaches to improve the accuracy of signal classification in an MI BCI system, particularly in discriminating different center-out movements of the same limb. Research includes the design of experimental BCI paradigms, signal acquisition, signal processing, and signal classification. The objectives will be achieved by applying information entropy and machine learning to increase classification accuracy. In addition, the experimental paradigm is modified by integrating kinesthetic vibrotactile stimulation to increase classification accuracy and analyze the effects on interaction workload. The research shows the value of vibrotactile guidance in improving the MI formation process and demonstrates that vibrotactile guidance improves MI extraction and subtly influences neurophysiological markers. Analysis of various time-frequency representations (TFRs) and their Shannon and Rényi short-term entropy measures for MI control signals in EEG data indicates that TFR-based entropy features lead to higher classification accuracies compared to regular amplitude features in the detection of MI. Finally, the evaluation of six classifiers, including commonly utilized and newly applied ones, showed that the ResNet-based classifier outperformed the others in detecting the different directions of MI. It also confirmed that preprocessing the data with low-frequency signal features improved classification accuracy, with vibrotactile guidance showing a significant influence, especially for simpler classifiers.

Keywords: kinesthetic vibrotactile guidance, motor imagery, brain-computer interface, information entropy, time-frequency representations, machine learning

PROŠIRENI SAŽETAK

Motorička predodžba (*eng.* motor imagery - MI) je kognitivni proces zamišljanja izvedbe motoričkog zadatka bez korištenja stvarnih pokreta mišića. Često se koristi u rehabilitaciji i u asistivnim tehnologijama za kontrolu sučelja mozak-računalo (*eng.* brain-computer interface - BCI).

Ova je disertacija usmjerena na predlaganje primjene novih pristupa s ciljem povećanja točnosti klasifikacije elektroencefalografskih (*eng.* electroencephalographic - EEG) signala pri korištenju MI BCI sustava za klasifikaciju različitih ortogonalnih kretnji istog ekstremiteta od središta prema vani. Proces istraživanja disertacije sastoji se od dizajna paradigme BCI eksperimenta te prikupljanja, obrade i klasifikacije signala. Prikupljeni signali sadrže različite namjere interakcije (različiti smjerovi i stanje mirovanja). Navedeni ciljevi postižu se korištenjem prilagođene predobrade, informacijske entropije i strojnog učenja u svrhu povećanja točnosti klasifikacije te promjenom eksperimentalne paradigme dodavanjem kinestetičke vibrotaktilne stimulacije u svrhu povećanja točnosti klasifikacije i analize utjecaja na interakcijsko opterećenje.

Korištenje različitih vremensko-frekvencijskih prikaza (*eng.* time-frequency representation - TFR) i njihovih Rényijevih i Shannonovih kratkotrajnih entropija za otkrivanje upravljačkih signala temeljenih na MI EEG podacima, rezultira većom točnošću klasifikacije u usporedbi s uobičajeno korištenim amplitudnim značajkama, što ukazuje na poboljšanje u otkrivanju MI-a. Konkretno, kratkotrajna Shannonova entropija temeljena na ponovno dodijeljenom pseudo Wigner-Ville TFR-u (*eng.* Reassigned pseudo Wigner-Ville TFR) postigla je točnost do 99,87% pri otkrivanju MI-a.

Evaluacija šest različitih klasifikatora, uključujući uobičajeno korištene linearne diskriminantne analize (*eng.* linear discriminant analysis - LDA), stroj potpornih vektora (*eng.* support vector machine - SVM), slučajne šume (*eng.* random forest - RF) i tri klasifikatora

iz obitelji konvolucijskih neuronskih mreža (*eng.* convolutional neural networks - CNN), koji su po prvi puta primijenjeni u ovom kontekstu, pokazuje da CNN koji se temelji na ResNet-u značajno nadmašuje konkurentske klasifikatore na podacima koji sadrže statične vizualne znakove, dinamičko vizualno navođenje ili kombinaciju dinamičkog vizualnog navođenja i dinamičkog vibrotaktilnog navođenja pri otkrivanju različitih smjerova MI-a. Konkretno, algoritam ResNet-101 postiže točnost do 72,30% pri klasificiranju različitih MI kretnji istog ekstremiteta. Studija također otkriva da predobrada podataka pomoću značajki niskofrekventnih signala daje veću točnost klasifikacije, te da vibrotaktilno vođenje značajno utječe na točnost klasifikacije, posebno za jednostavnije klasifikatore.

Ključne riječi: kinestetičko vibrotaktilno navođenje, motorička predodžba, sučelje mozak-računalo, informacijska entropija, vremensko-frekvencijski prikazi, strojno učenje

CONTENTS

Chapter 1

INTRODUCTION

The loss or impairment of motor functions due to illness or accident has a significant impact on the lives of those affected. Therefore, specialists in various disciplines are striving to explore methods by which these lost abilities can be restored or replaced. In doing so, it is crucial to take a close look at the sensorimotor control loop in its entirety, which includes both efferent and afferent processes. In the context of natural movement, efferent feed-forward (consisting of movement execution) and afferent feedback components (consisting of haptic data, proprioception, visual input, and more) are inseparable [1]. Sensory inputs constantly adjust and refine motor actions during movement performance. In recent years, brain-computer interfaces (BCIs) have gained significant attention due to their potential applications in motor rehabilitation and assistive device control. One of the key aspects of BCIs is their ability to interpret motor imagery (MI) signals from the user's brain, allowing them to perform tasks without having to move. The combination of MI, a cognitive process in which a person imagines executing a motor task without actually executing the movement, and afferent feedback has been shown to be effective [2, 3, 4]. Moreover, effective interpretation, detection, and classification of MI signals can greatly enhance the utility and performance of BCIs, leading to better user experiences and outcomes. Electroencephalography (EEG) has emerged as a popular method for capturing MI signals due to its high temporal resolution, relatively low cost, and ease of use. The following chapter delves into the principles, methods, and applications of EEG and provides a comprehensive understanding of this technology and its role in advancing BCI research for motor rehabilitation and assistive device control.

1.1. Electroencephalography

EEG is widely employed as a non-invasive technique for capturing the electrical activity of the brain, known as an electrogram (shown in Fig. 1.1). The electrical activity of the brain is measured by EEG through the detection of voltage fluctuations that arise from the synchronized activity of neurons in the cerebral cortex [5]. The communication between neurons occurs via synaptic transmissions, which in turn generate electrical currents. These currents, when combined, give rise to local field potentials that can be effectively detected by electrodes placed on the scalp. As a result, the EEG signals obtained reflect the overall brain activity, providing a representation of the collective output originating from millions of neurons [6].

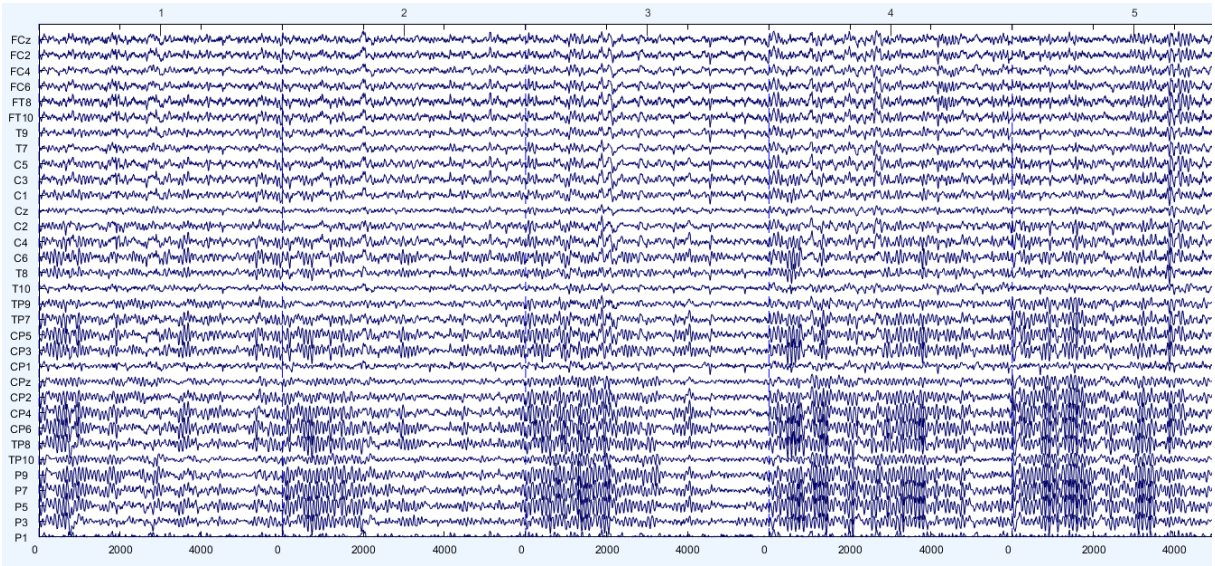


Figure 1.1: Example of an EEG electrogram, from the EEGLAB Matlab toolbox [7].

The process of data recording in EEG involves the placement of electrodes on the scalp prior to the commencement of the recording. The type and positioning of these electrodes significantly influence the quality and reliability of the EEG recordings. Two main types of electrodes are commonly utilized in EEG recordings: wet electrodes and dry electrodes. Wet electrodes require the application of a conductive gel (or saline solution for some specific electrodes) to establish a reliable connection between the electrode and the scalp, whereas dry electrodes do not require additional substances. Each type of electrode offers distinct advantages and disadvantages. Wet electrodes generally offer superior signal quality, however, they require more preparation time and maintenance. On the other

hand, dry electrodes are easier to apply, but they may be associated with lower signal quality. The placement of electrodes typically follows the guidelines of the International 10-20 system, which ensures consistent positioning across different individuals and studies [5]. This standardized system divides the scalp into various regions, and the electrodes are positioned at specific percentages of the distance between anatomical landmarks such as the nasion and inion, as well as the preauricular points [5].

The brain's electrical activity recorded and observed by EEG originates from neurons within the underlying brain tissue. This implies that the recorded EEG signals are influenced by various factors, including the orientation and proximity of the scalp electrodes to the source of neuronal activity. Additionally, the recorded values can be distorted by intervening tissues and bones. Consequently, the EEG signal does not equally represent the activity of all neurons, rather, it predominantly reflects the synchronous activity of neurons closer to the electrodes. Deeper brain structures (e.g., the brain stem or hippocampus) do not directly contribute to the EEG signal [5, 6]. It is important to consider these factors when interpreting EEG recordings and understanding the limitations of this non-invasive technique.

In addition to EEG, there are various other electrophysiological monitoring methods used to investigate brain activity. One such method is electrocorticography (ECoG), which shares similarities with EEG but is invasive as it requires the surgical placement of electrodes directly on the surface of the brain. By positioning the electrodes closer to the neural tissue, ECoG offers the advantage of obtaining more precise readings of electrical activity in deeper brain structures. The placement of electrodes on the cortical surface allows for a more direct and localized measurement of neuronal activity, enabling researchers to study specific brain regions with greater accuracy.

Functional magnetic resonance imaging (fMRI) is a non-invasive neuroimaging technique that utilizes magnetic fields to measure changes in blood oxygenation levels, enabling the identification of brain regions involved in specific cognitive processes and tasks. However, compared to EEG, fMRI has lower temporal resolution, a higher cost, and limited experimental flexibility due to the constraints of the MRI scanner environment, which can limit the range of tasks and experiments that can be conducted [5].

Functional near-infrared spectroscopy (fNIRS) is a non-invasive neuroimaging technique that measures changes in blood oxygenation levels in the brain. It utilizes near-

infrared light to penetrate the scalp and skull and reach the cortical surface, where it interacts with hemoglobin in the blood. fNIRS relies on the principle of neurovascular coupling, which states that changes in neural activity lead to alterations in cerebral blood flow and oxygenation. By measuring these hemodynamic changes, fNIRS provides insights into brain function and activation patterns. fNIRS offers several advantages as a neuroimaging technique. It is portable, non-invasive, and well-suited for use in natural environments or during real-time interactions. fNIRS can complement other neuroimaging methods like EEG and fMRI by providing additional spatial and temporal information. However, it is important to note that fNIRS has limitations. The depth of penetration is limited, primarily capturing signals from the outer cortical layers [5].

Electromyography (EMG) is a technique used to record the electrical activity generated by skeletal muscles. It is primarily employed to assess muscle function, diagnose neuromuscular disorders, and guide rehabilitation interventions. Unlike EEG, EMG specifically measures electrical activity from muscles and is not suitable for assessing brain activity directly, but it can still be utilized in BCIs. EMG recordings can be influenced by factors such as muscle fatigue, crosstalk between adjacent muscles, and the quality of the skin-electrode interface. EMG is also more susceptible to movement artifacts and electrical interference compared to EEG [5].

Magnetoencephalography (MEG) is a non-invasive method for measuring the magnetic fields produced by neuronal activity in the brain. It offers high temporal resolution and improved spatial accuracy compared to EEG, making it particularly suitable for studying the dynamics of cognitive processes. MEG's main downside compared to EEG is the considerably higher cost associated with the equipment and its maintenance. MEG systems also require specialized, magnetically shielded rooms, making them less accessible and less portable than EEG systems. Additionally, MEG is less sensitive to radial sources of brain activity, which can result in the underestimation of certain brain signals compared to EEG [5].

A healthy human EEG exhibits distinct patterns of activity that correspond to the individual's level of wakefulness. These activity patterns are characterized by specific frequency ranges, typically ranging from 0.2 to 100 Hz. These frequencies are further categorized into different bands, including alpha ($8 - 13$ Hz), mu ($8 - 12$ Hz), beta ($15 - 32$ Hz), delta ($0.2 - 5$ Hz), theta ($4 - 7$ Hz), and gamma ($30 - 100$ Hz) [1].

Alpha waves predominantly occur in the posterior regions of the brain and are typically observed during states of resting wakeful state. They are most visible over the occipital and parietal lobes of the brain (depicted in Fig. 1.2). On the other hand, mu waves are present in the motor cortex and sensory cortex, also known as the sensorimotor cortex. While alpha and mu waves share similar frequency ranges, their spatial distribution within the brain differs [6].

Beta waves, characterized by higher frequencies, become more prominent in frontal areas and other regions during intense mental activity. For instance, when a person in a relaxed state opens their eyes, alpha activity decreases while beta activity increases, reflecting the transition to a more active cognitive state [6].

Theta and delta waves are typically not observed during resting wakefulness but can increase in amplitude during certain mental activities, such as specific stages of sleep, deep meditation, or even certain mental tasks [8]. These lower frequency waves are associated with slower brain activity [6].

In terms of EEG amplitude, it can vary between 20 and $100\mu V$. To enhance the signal quality and distinguish it from background noise, amplification of the EEG signals is essential. This amplification helps improve the signal-to-noise (SNR) ratio, making the signals more suitable for further analysis. Additionally, various filtering techniques (such as band-pass filtering, component analysis, and spatial filtering) are employed to remove unwanted artifacts and noise from the recorded EEG signals. These artifacts can include electrical line noise, muscle activity, and eye movement artifacts. Filtering ensures that the signals of interest are preserved while minimizing the influence of unwanted noise, allowing for more accurate interpretation and analysis of the EEG data.

EEG's ability to identify abnormal brain activity (such as sharp waves, spikes, and discharges) makes EEG a valuable tool for medical diagnosis in the field of epilepsy. Moreover, EEG is extensively utilized in various other areas of medical diagnosis, including the assessment of sleep disorders, the evaluation of coma and encephalopathies, and the assessment of brain death [6]. In addition to its clinical applications, EEG is also widely used in cognitive sciences and serves as one of the primary methods for gathering control signals in BCIs. Despite its limited spatial resolution, EEG remains an indispensable tool for research and diagnosis due to its mobility and remarkable temporal resolution in the millisecond range. This high temporal resolution is unmatched by many other

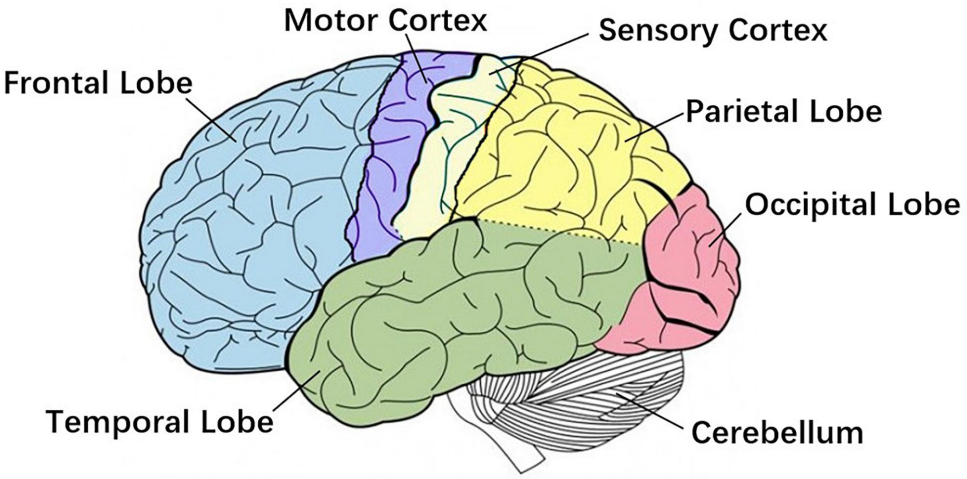


Figure 1.2: The lobes of the brain, with motor cortex (part of the frontal lobe) and sensory cortex (part of the parietal lobe) emphasized [9].

measuring techniques, allowing researchers to capture and analyze the rapid dynamics of brain activity with great precision. Therefore, EEG continues to be one of the preferred choices for studying brain function and its abnormalities, enabling advancements in both clinical and research domains [6].

EEG derivatives encompass various analysis techniques that provide deeper insights into brain function and cognitive processes. One prominent derivative is event-related potentials (ERPs), which involve averaging EEG responses time-locked to specific sensory, cognitive, or motor events. ERPs are widely utilized in neuroscience, BCIs, cognitive science, cognitive psychology, and psychophysiological research to investigate the brain's response to complex stimuli. This technique focuses on the time-domain analysis of EEG signals, examining the amplitude, latency, and shape of the recorded potentials [5]. To calculate ERPs, multiple EEG epochs related to the event of interest are averaged together, which improves the SNR and reveals underlying neural processes. This approach allows researchers to identify specific components and analyze their characteristics, the brain's cognitive processing, and sensory perception in response to different stimuli.

Frequency-domain analysis is another vital aspect of EEG analysis. It involves transforming the EEG signals from the time domain to the frequency domain, enabling the identification of distinct frequency bands associated with different brain states and cognitive processes [5, 10]. Fourier analysis is a commonly used method for this transformation, allowing researchers to extract the power spectral density (PSD) of the EEG signals. The

PSD represents the distribution of signal power across different frequencies and provides valuable information about the brain's oscillatory activity. Furthermore, time-frequency analysis combines the strengths of both time-domain and frequency-domain analysis, offering a more comprehensive understanding of EEG signals. Techniques such as the short-time Fourier transform (STFT), spectrogram, and wavelet transform allow researchers to explore the time-varying nature of frequency components in the EEG data. This dynamic analysis provides valuable insights into the temporal dynamics of neural processes, revealing changes in frequency components over time.

By employing these analysis methods, researchers can unravel the intricate dynamics of the brain's electrical activity and gain a deeper understanding of the underlying cognitive processes. EEG derivatives, including ERPs, frequency-domain analysis, and time-frequency analysis, contribute significantly to advancing collective knowledge of brain function and cognitive neuroscience.

1.2. Brain-computer Interfaces

BCIs are an emerging field of interdisciplinary research that aims to establish direct communication between the human brain and external devices, such as computers, assistive robotics, or virtual environments [5]. BCIs have the potential to significantly impact various domains, including medical rehabilitation, assistive technologies for individuals with disabilities, gaming, and human-computer interaction (HCI). This chapter provides an overview of the principles, methods, and applications of BCIs, focusing on their role in motor rehabilitation and assistive device control. An example of a BCI system can be seen in Fig. 1.3.

BCIs, an exciting and interdisciplinary field, seek to establish a direct communication channel between the human brain and external devices such as computers, assistive robotics, and virtual environments. With their immense potential, BCIs can revolutionize several domains, including medical rehabilitation, assistive technologies for individuals with disabilities, gaming, and human-computer interaction (HCI). This section provides a comprehensive overview of the basic principles, methods, and applications of BCIs, with particular emphasis on their role in motor rehabilitation and assistive device control.

Fig. 1.3 illustrates a specific example of a BCI system, showcasing the components

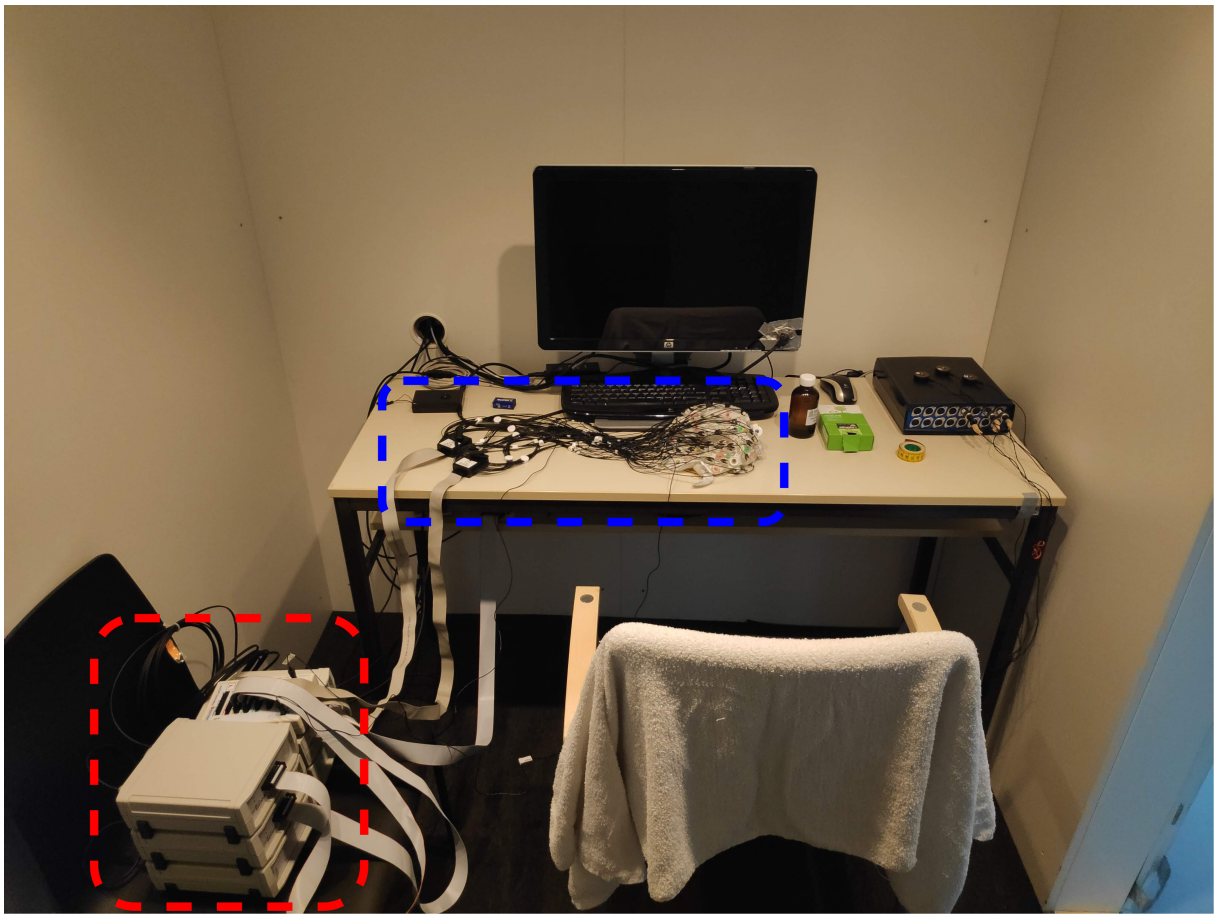


Figure 1.3: An example of a BCI system utilizing EEG. The amplifiers are marked with a red rectangle, and the active electrodes and cap are marked with a blue rectangle. During the experiment, the subject sits on the chair in front of the screen, and the screen displays cues and instructions about the experiment.

and interactions involved in this technology. The BCI system acts as an intermediary that enables the bidirectional flow of information between the user's brain and the external device (e.g., a computer screen or robotic arm). Understanding the underlying principles and techniques of BCIs is essential to unlocking their full potential. By achieving advancements in brain signal acquisition, processing, and interpretation, researchers can develop effective strategies to harness the brain's electrical activity for BCIs.

BCIs can be categorized into two broad types: invasive BCIs and non-invasive BCIs. In invasive BCIs, such as ECoG BCIs, electrodes are implanted directly into the brain tissue. This invasive approach allows for high spatial resolution and excellent signal quality, making it suitable for precise neural recordings. However, it also carries risks such as the potential for infection, tissue damage, and problems with long-term stability. The invasiveness of this method requires careful testing and monitoring to ensure the safety

and efficacy of implanted devices. Non-invasive BCIs, on the other hand, use external sensors to measure brain activity without the need for surgical intervention. Common non-invasive techniques can utilize EEG, fNIRS, or MEG. Although non-invasive BCIs generally have lower spatial resolution and signal quality compared with invasive approaches, they offer significant advantages in terms of accessibility, portability, and safety. Non-invasive BCIs are more user-friendly and can be applied in a wider range of settings without the risks associated with invasive procedures. The choice between invasive and non-invasive BCIs depends on several factors, including the specific application, the desired level of precision, the individual's condition, and the trade-off between precision and safety. Researchers and clinicians carefully consider these factors when selecting the appropriate BCI approach for a particular scenario. Both invasive and non-invasive BCIs have made remarkable advancements, and ongoing research continues to explore ways to improve their performance, reliability, and ease of use.

BCIs employ various paradigms to decode and interpret brain activity to enable communication and control of external devices. Some examples of such paradigms are MI, ERPs, steady-state visually evoked potentials (SSVEPs), slow cortical potentials (SCPs), and local field potentials (LFPs). MI is a popular paradigm in the field of BCIs that allows users to control devices or perform actions by mentally rehearsing or imagining the execution of certain motor actions without actually physically moving [11]. For example, users can imagine moving their hand, arm, or leg or even performing complex tasks such as playing an instrument or typing on a keyboard. During MI tasks, the brain generates neural activity patterns that are similar to those observed during actual movement execution. One of the advantages of BCIs based on MI is their versatility and potential applicability to a wide range of individuals, including those with motor disabilities or paralysis. MI allows users to interact with the environment and control external devices through their thoughts alone, providing them with a means of communication, mobility, and independence. Another commonly studied paradigm type is the P300 ERP [12], which is a positive deflection in the EEG waveform that occurs about 300 ms after the presentation of a rare or target stimulus in a stimulus stream. The P300 ERP is associated with cognitive processes related to attention, memory, and decision-making [12]. During a P300-based BCI task, the user is presented with a series of stimuli, and his or her task is to focus attention on the target stimulus. The target stimulus is usually different

from the non-target stimuli, and its presentation triggers a P300 response. By detecting and analyzing the P300 response in the EEG signals, the BCI system can infer the user's intention or selection of the target stimulus. It is known that P300 can be successfully applied to both visual and auditory stimuli [13, 14, 15]. SSVEPs are rhythmic oscillations in the EEG that occur at the same frequency as a visual stimulus that flickers at a constant rate. By presenting visual stimuli with different flicker frequencies, users can direct their attention to a particular stimulus, and the corresponding SSVEP can be detected and used for device control [16]. SCPs are gradual shifts in brain electrical activity that can be voluntarily modulated. Users can learn to self-regulate their SCPs through techniques such as biofeedback and operant conditioning. By differentially controlling their SCPs, users can create distinct patterns in the EEG that can be used as control signals for BCIs [17]. LFPs are low-frequency electrical signals that arise from the collective activity of neurons within a specific brain region. LFP-based BCIs typically involve electrodes implanted directly into brain tissue, which allows recording and analysis of neuronal activity with finer spatial resolution. These signals can provide valuable insights into the local circuitry and dynamics of the targeted brain area [18].

These paradigms offer different approaches to interacting with BCIs that accommodate different user abilities and preferences. By understanding and leveraging the neural mechanisms underlying these paradigms, researchers and engineers can develop effective BCI systems that allow individuals to interact with the outside world using their brain activity.

MI-based BCIs have been used in various fields, including neurorehabilitation, assistive technology, and virtual reality. In neurorehabilitation, MI can be utilized as a therapeutic tool for motor recovery by helping people with motor impairments engage in MI tasks to stimulate brain plasticity and facilitate motor rehabilitation through repetitive practice, mental rehearsal, and real-time feedback. In assistive technology, MI-based BCIs' goal would be to enable individuals with physical disabilities to efficiently control prosthetic devices (e.g., limbs), robotic devices, or computer interfaces, thereby restoring their ability to interact with the world. BCIs can enable people with severe motor disabilities to interact with their environment and regain a degree of independence. For example, BCIs can be used to control spellers and communication devices [14, 19] and potentially control wheelchairs or prosthetic limbs by translating the user's MI signals or other cognitive

states into corresponding commands [11].

Despite the significant advancements in BCI research, several challenges remain to be addressed. These include improving the reliability and accuracy of BCI systems, reducing the time and effort required for user training, and developing more intuitive and user-friendly interfaces.

1.3. Sensorimotor Rhythms, Motor Imagery, and Somatosensory Loop

One of the most important phenomena observed in EEG recordings (and other neuroimaging recording techniques) is that the neurophysiological rhythmic activities observed and recorded over the sensorimotor cortex are modulated by actual movement execution, movement intention, or MI. The modulation manifests as amplitude decreases in the alpha, mu, and beta frequency bands, also called event-related desynchronization (ERD) [1]. Mu rhythms, also known as Rolandic rhythms or sensorimotor oscillations, are typically observed in the frequency range form 8 to 12 Hz. They are most pronounced over the sensorimotor cortex, which encompasses both primary motor and primary somatosensory areas. Mu rhythms are associated with various aspects of motor function, including motor planning, execution, and resting state. Interestingly, mu rhythms are not only attenuated or desynchronized during actual motor execution, but also during MI, making them particularly relevant for BCI applications. Beta rhythms, with a frequency range of $15 - 32$ Hz, are another essential component of sensorimotor rhythms (SMRs) [1]. These rhythms are involved in various cognitive processes related to motor function, such as motor anticipation, coordination, and maintenance of the current motor state. Beta rhythms are also known to be modulated during motor tasks and MI. They show ERD during preparation and execution of movements and event-related synchronization (ERS) during completion of movements or in the post-movement period. ERS is an amplitude increase in the gamma frequency band (>30 Hz) [20]. Examples of the cortical distribution of ERD and ERS during left and right-hand MI tasks are shown in Fig. 1.4.

ERD and ERS are rhythmic activities that can also be interpreted within SMRs [21]. Kobler et al. [22] demonstrated that modulations of the EEG delta band ($0.2-5$ Hz) also

contain information about the onset and direction of arm movements, involving two separate cortical networks. Direction information was primarily encoded in parieto-occipital areas synchronized with the direction cue, whereas arm movement onset was reflected in central sensorimotor areas through the movement-related cortical potentials (MRCPs)[1]. MRCPs are neural-specific patterns associated with self-directed initiation of movement [23]. The pattern is characterized by a progressive increase in negativity that begins about 1.5 s before movement initiation and reaches a peak of negativity near movement initiation. MRCPs also encode properties of the movement, such as speed, applied force [1, 24, 25], or directional information [22].

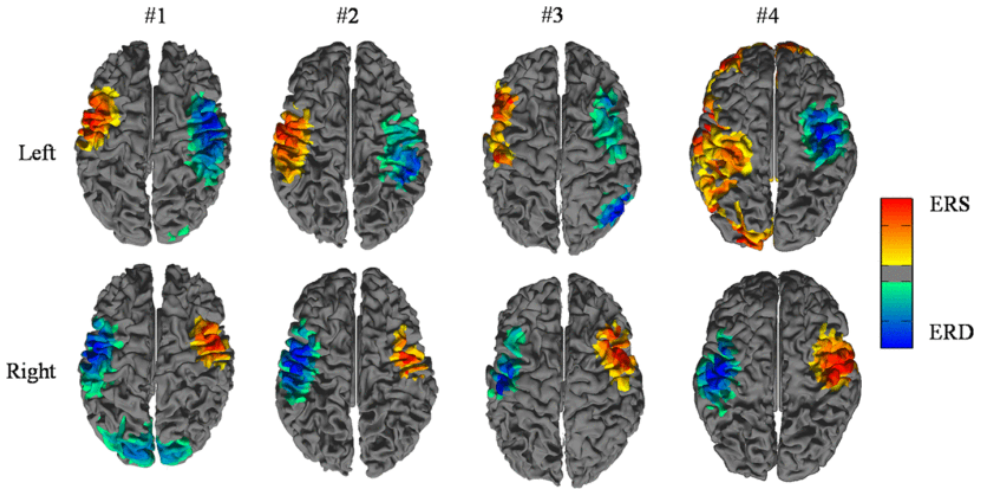


Figure 1.4: Cortical distribution of mu ERD and ERS during cursor movement for four subjects [26] ©2008 IEEE.

SMR-based BCI, which is the prerequisite for brain control in such systems, can capture motor intentions, or MI. Several studies have demonstrated that individuals can control the amplitude of SMR via MI [27, 28, 29, 30, 31]. In several experiments, subjects were able to gain control in both two-dimensional (2D) and three-dimensional (3D) spaces [28, 31]. The sources of sensorimotor rhythms triggered by movements or imagined movements of various body parts have been localized in the primary sensorimotor cortex in a somatotopic manner [32]. To date, BCIs with MI provide the highest level of control in terms of degrees-of-freedom (DOF) among all other signal components [33] (e.g., ERP or SCP) [34].

It is important to note that SMR can be influenced by various external factors, such as attention, cognitive load, and sensory input. For example, the presence of dynamic visual or tactile guidance or feedback during a motor task can modulate SMR activity, potentially

altering the effectiveness of BCI systems that rely on these rhythms (as opposed to BCI systems in which only visual cues are present) [1, 35]. Understanding the effects of these external factors on SMRs is critical to optimizing BCI performance and ensuring the success of SMR-based experiments.

It is necessary to consider feed-forward processes, which entail the execution of movements, and feedback processes, such as haptic information, proprioception, and visual information, as interrelated in natural locomotor activities [1], because they provide essential information to the nervous system about body position, movement, and interaction with the environment [1, 35]. Proprioception, the sense of body position and movement, relies on specialized receptors within muscles, tendons, and joints that convey information about changes in muscle length and tension [36]. Somatosensory input encompasses a range of sensations such as touch, pressure, vibration, temperature, and pain that are sensed by specialized receptors in the skin and deeper tissues [37]. These sensory inputs continuously shape and refine motor actions during execution, forming an intricate sensory feedback loop (Fig. 1.5) that enables precise and coordinated movements. Adjustments and refinements to the movement sequences occur during execution and are influenced by the sensory inputs. The integration of MI, in which participants imagine performing actions without physically carrying them out, with afferent feedback has demonstrated the induction of plasticity in motor cortex [2, 3, 38]. The study by Mrachacz-Kersting et al. [38] found a strong correlation between electrical stimulation-induced afferent influx and EEG motor-related patterns in the amplitude of low-frequency bands, suggesting a simultaneous arrival at the somatosensory areas. Exclusion of somatosensory feedback from natural movement results in reduced motor control during reaching movements in which haptic perception has been artificially eliminated [39], and during wrist movements in which proprioception has been artificially disrupted [40]. Especially in the context of upper limb prostheses for amputees, many users express the need for tactile feedback [41, 42, 43, 44] so that the interaction feels more like a natural movement [1].

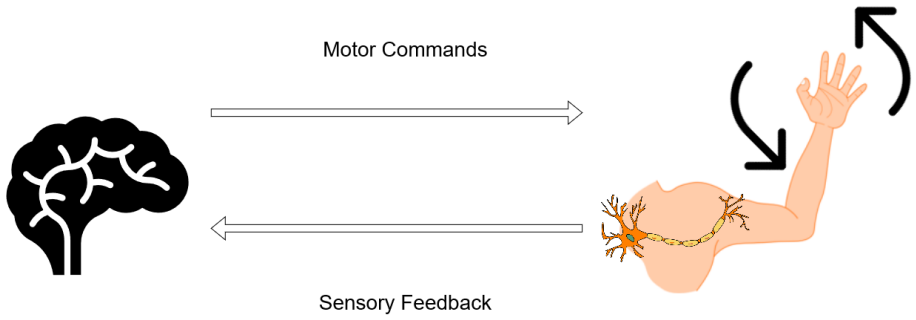


Figure 1.5: Illustration of a somatosensory loop.

1.4. Hypotheses and Contributions

Previous research, described in detail in Chapter 2, has explored to some degree the potential benefits of artificial somatosensory input for MI performance in applications where MI is used as a control tool for end-effectors. Various approaches to providing this artificial input have been explored. In the context of rehabilitation, research has shown promising results regarding the effects of passive movement feedback on motor outcomes and BCI performance. Some studies have even found a subjective preference for vibrotactile feedback over visual feedback, suggesting that different feedback modalities may influence users' experiences and outcomes. It has come to the attention of the author of this dissertation that there is a lack of solutions in which the potential influence of vibrotactile input on directional classification and MI detection (classification of MI against baseline or rest) from EEG has been investigated.

There are numerous approaches to processing MI signals, and information entropy has shown promise in this area. As described in Chapter 2, a variety of entropy applications have been explored in EEG SMR studies for different purposes. For example, resting state EEG spectral entropy has been used as a biomarker for predicting SMR BCI performance, and Shannon entropy has been used for detecting the intentional non-control (INC) state to improve the usability of BCI by reducing unintended commands during SMR BCI operations [33]. Other studies have focused on utilizing entropy for motion detection and prediction and improving MRCP detection. Researchers have also combined different entropy features, such as amplitude-based Shannon entropy, phase-based Shannon entropy, wavelet entropy, and sample entropy, to achieve moderately high classification accuracy

for MI detection. Despite the successful application of various entropy measures in previous studies, short-term information entropy based on time-frequency representations (TFRs) interpreted as probability density functions has not been extensively studied [33]. This approach could potentially provide new insights and opportunities to improve the processing and interpretation of MI signals.

Classification accuracy is a crucial comparative metric in studies using MI BCIs. Related work has shown that it is easier to achieve higher classification accuracy for different-limb different MIs than for same-limb different MIs. When studying the classification of MI with the same limbs, simpler machine learning (ML) methods are often employed, resulting in a wide range of accuracy levels, some at questionably useful chance level, while others achieve significantly higher performance. To date, no research has applied recently developed ML techniques with more complex architectures, such as deep convolutional neural networks (CNNs), to classify MI tasks with the same limbs [27]. The goal of this part of research is to fill the current research gap and augment the domain of ML-based EEG analysis. This will be achieved by proposing and demonstrating a novel processing and classification pipeline, including specialized deep CNNs tasked with classifying different MIs of the same limb. The ultimate aim is to surpass the current state-of-the-art in classification accuracy described in the Chapter 2.

Based on the observed research gaps and shortcomings of representative paradigm designs, classification accuracy achieved during MI detection, and classification accuracy achieved during classification of different MIs of the same-limb, the research hypotheses were defined:

1. Vibrotactile stimulation can be utilized in an MI experiment to increase the classification accuracy of interaction intentions (classification of different MIs and MI against rest state) and decrease the interaction workload while using the BCI system;

2. The quantity of classifiable MI information (entropy output) changes during the period of MI in comparison to a rest state (baseline). As a result, classification accuracy will increase when classifying features obtained from entropy compared to classification accuracy when using commonly utilized amplitude features, thus improving the ability to detect MI;

3. By applying the adapted classification pipeline to the original curated dataset as well as to the publicly available dataset, the detectability of MI and the successful detection of different interaction intentions can be further improved.

Based on the above-stated hypotheses, the scientific contributions of this research are also defined:

1. A new method that uses kinesthetic vibrotactile guidance in an MI BCI experiment in order to increase the classification accuracy of interaction intentions;

2. Entropy-based methods that improve the detection accuracy of interaction intentions in the MI BCI experiment;

3. Original dataset of MI BCI data, involving kinesthetic vibrotactile guidance;

4. The introduction of a unique application for particular deep CNNs is proposed, with the objective of enhancing the classification accuracy of interaction intentions in same-limb MI tasks.

1.5. Research Methodology

The research was conducted in five phases, which are closely related to the realization of the scientific contributions of this research:

– The first phase of the research included a detailed review of the scientific literature relevant to the research area. This review includes scientific papers that address the effects of artificial somatosensory input for BCI experiments that utilize MI for directional decoding, specifically center-out movements. The review of the literature also included scientific papers dealing with the use of the information entropy measure for MI detection. Particular attention was paid to the field of information entropy obtained from TFRs when applied to MI data, as well as the use of TFRs on EEG data in general. In addition to the mentioned areas, this review also focused on the current state-of-the-art in classification methods and their performance in the field related to the subject area in order to find the most appropriate classifier for the characteristics of the datasets used.

– The second phase consisted of designing the paradigm, setting up the experiment, and collecting the data. Data were recorded from 15 subjects. Participants sat in front of a computer screen and were instructed to perform the MI center-out task of the same limb. The task was guided by visual guidance displayed on the screen and simultaneous vibrotactile moving sensation, or by visual guidance alone. Participants were adequately informed about the experiment and underwent an initial familiarization procedure at the beginning of the experiment. This preparation was to ensure that they could optimally perform the MI tasks. Participants completed a sufficient number of runs for subsequent processing and classification purposes. In addition to the main part of the experiment, additional runs were recorded with each participant for controlled eye artifacts and rest (which were also used during processing and classification). At the conclusion of the experiment, participants were asked to complete a questionnaire. This questionnaire was designed to gather insights about their personal experiences during the course of the experiment. In this phase, curation of the new dataset laid the foundations for the first two scientific contributions, and the third scientific contribution was achieved.

– In the third phase, the newly curated data were preprocessed and prepared for classification and further processing. The processing consisted of downsampling the data, filtering to the appropriate frequency bands, removing artifacts and bad trials, and calculating the TFR and spectral power characteristics of the data. In this phase, the neurophysiological background of the newly obtained data was also investigated and interpreted (movement-related cortical potentials, event-related synchronization and desynchronization, amplitude potentials, and power of the features). In this phase, the results of the participants' questionnaires and their perceived experience of the experiment were analyzed and interpreted. A basic classification was carried out in order to compare the different features and their effects on MI detection and MI classification. Three different aspects were classified in the study. First, the classification of MI for different movements was carried out for each condition (with and without vibrotactile guidance) separately. This evaluation was used to assess the effect of vibrotactile guidance on direction discrimination. Subsequently, the MI period was compared to the rest state (or baseline) for each condition separately.

The goal was to investigate the effect of vibrotactile guidance on the detection of the MI state. Finally, the two conditions were classified against each other in order to determine the detectability of vibrotactile guidance in such MI task. The first scientific contribution of this research was realized in this phase.

– In the fourth phase, information entropy features and the performance of the classification accuracy of various TFRs used with different entropy features were analyzed. The main objective of this phase was to explore the effectiveness of the short-term entropy based on different TFRs for the detection of MI, thereby investigating their effectiveness for the detection of the INC state. With the newly obtained short-term entropy features, the classification of MI of different movements and the classification of MI compared to the rest state were performed separately for each condition. When all the classification results of this phase were available and the validity of the proposed methods was confirmed, another dataset (suitable and publicly available) was introduced to test the proposed processing pipeline and methods. The second scientific contribution was made in this phase.

– In the final fifth phase of the dissertation, a comprehensive benchmarking and performance analysis of six different classifiers – three commonly utilized classifiers and three newly deployed CNN types – was performed using EEG MI datasets. The primary objective was to evaluate the effectiveness of the classifiers when applied to different MI guidance modalities, including static visual cues, dynamic visual guidance, and a combination of dynamic visual guidance and vibrotactile guidance. An additional aspect of the study was to evaluate the impact of data preprocessing, particularly passband filtering, on the performance of these classifiers. This phase aims to find a novel application of specific deep CNNs to improve same-limb MI classification accuracy while outperforming traditional standards. The effects of varying guidance techniques and different data preprocessing approaches on classification accuracy were statistically scrutinized for all methods studied. As a result of the research efforts in this phase, the fourth scientific contribution was made and, the first scientific contribution was statistically corroborated.

The research conducted in this dissertation can be divided into three distinct but interrelated research segments, each supported by the author's published work. The

segments are named for ease of distinction and navigation throughout the later chapters of the dissertation. The first segment is named *KGU research* (the name suggests that it comes from a study in which **K**inesthetic **GU**idance was integrated into the paradigm) and deals with the second and third phases described above. The second segment is named *Entropy research* and deals with the fourth phase. Lastly, the third segment is named *Classification methods comparison research* and is mainly included in the fifth phase.

In this dissertation, certain parts of the research are a reproduction of the author's previous publications, namely "Directional Decoding From EEG in a Center-Out MI Task With Visual and Vibrotactile Guidance" [1], "Detection of motor imagery based on short-term entropy of time–frequency representations" [33], and "Motor Imagery Classification Based on EEG Sensing with Visual and Vibrotactile Guidance" [27]. These articles are openly accessible and distributed under the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/), which allows for the reproduction and adaptation of the material with appropriate credit.

1.6. Structure of the Doctoral Dissertation

This dissertation is divided into seven chapters. In this first chapter, a detailed background encompassing EEG, SMR, MI, and the somatosensory loop is thoroughly explored, providing a robust foundation for the content of the following chapters. In addition, this opening chapter clearly lays out the assumptions and scientific contributions that drive this research and sets the direction and expectations for this study.

In the following Chapter 2, a comprehensive review and critical assessment of previous research related to MI, and in particular, MI with somatosensory input, is presented. This review also includes an examination of the role of information entropy in MI and BCI, as well as a survey of the current state-of-the-art MI BCI classification techniques. This review forms the basis for the innovative work described in the following chapters.

Presenting a unique methodology, Chapter 3 outlines the process behind the curation of an original dataset, which includes the development of a new experimental paradigm, finding participants, the data collection phase, and the questionnaire regarding the subject's perceived experience of the experiment. This chapter also provides important details about the second dataset used in this research, which was not curated within this study

but is publicly available. Both datasets represent different center-out MIs of the same limb, offering a comparative perspective.

In Chapter 4, details of the signal processing and preparation methods applied to both datasets for further analysis are presented. This chapter also examines the comparison of different TFRs and their Rényi and Shannon short-term entropies for the detection and classification of different MIs of the same limb to demonstrate the value of these measures for improving MI detection accuracy.

Chapter 5 presents and describes all classification methods and features used in the research conducted as part of this dissertation and provides insights into the suitability of different classification methods for various application contexts in MI BCI.

The neurophysiological background of the newly obtained data, the influence of the additional vibrotactile guidance in the same-limb different MIs experiment, and the subjects' perceived experiences with this new experiment are described in Chapter 6. In addition, the impact of utilizing short-term information entropy features for the detection of MI was investigated, and the evaluation of six different classification methods based on different types of features was presented. Finally, this chapter provides a discussion of all the findings and their implications.

Unveiling the concluding remarks, Chapter 7 brings together all the integral parts of this dissertation. It reflects upon the primary findings and highlights how they contribute to the MI BCI field. In particular, it addresses the impact of vibrotactile guidance, the value of information entropy, and the performance of the best classification methods. The chapter also addresses the broader implications of the study, shedding light on its potential applications and implications for assistive technologies. In addition, this conclusion presents a review of the limitations of the current work and potential directions for future research based on the outcomes of this study.

Chapter 2

REVIEW OF EXISTING STUDIES

This chapter provides a comprehensive analysis and synthesis of the relevant literature and research studies pertaining to the topics of somatosensory feedback, entropy analysis, and classification in the context of different MI tasks. This chapter serves as a critical examination of existing knowledge and aims to identify key findings, methods, theoretical frameworks, and gaps in the literature. By reviewing and evaluating a wide range of studies conducted in the field, this chapter aims to establish a strong foundation for current research and contribute to the advancement of knowledge in the field of different MIs of the same limb.

The study of motor execution and motor imagination was facilitated by the use of low-frequency EEG amplitude in both cue-based tasks [8, 45] and self-paced tasks [46, 47, 48]. According to the results of Mrachacz-Kersting et al. [4], positive effects on MRCP variability in stroke patients performing a motor execution task were produced by manual pressure stimulation. Additionally, an enhancement in cortical excitability was observed by Mrachacz-Kersting et al. [38] when healthy participants received either functional electrical stimulation or passive movement stimulation in response to the MRCP of imagined foot movements. Research by Kobler et al. [22] showed that directional information is encoded mainly in the low-frequency delta band and that discriminability between directions depends strongly on the alignment of the data.

For studies or applications that lack inherent somatosensory input, researchers often rely on replacing it with visual input. Visual input offers a number of possibilities, as it can process an extensive and diverse range of stimuli. This makes visual input an attractive

alternative for providing somatosensory input in such contexts. However, visual input is not the only modality that can replace somatosensory input [35].

The most common method of delivering non-invasive somatosensory input is vibrotactile stimulation [49, 50, 51, 52] in which mechanical vibrations are transmitted to the skin to produce tactile sensation. This modality has been used for various purposes, such as providing force feedback [51] or the transmission of encoded patterns with discrete properties [49, 50]. Electrotactile stimulation [38, 53, 54, 55, 56] is another commonly used modality in which electrical currents are applied to the skin to produce tactile sensations. Electrotactile stimulation has been used to transmit kinesthetic information for navigational purposes [53] force feedback applications, or encoded patterns with continuous properties in the tactile patterns [54]. Mechanotactile stimulation [57, 58]involves the application of physical pressure or mechanical displacement to the skin to elicit tactile sensations. This modality has been used primarily for force feedback applications. Passive movement [38, 59, 60, 61] refers to the physical movement of a body part or limb by an external source (e.g., an exoskeleton) without the active participation of the individual. Passive movements can be used to transmit kinesthetic information for proprioceptive purposes [60, 61]. By manipulating the movement of the body, individuals can gain a sense of spatial orientation and motion. Each of these modalities offers unique benefits and can be used for various somatosensory input applications. By understanding the specific properties and effects of these input modalities, researchers can develop more effective and immersive tactile interfaces for various purposes, including virtual reality, rehabilitation, and prosthetics [1, 35, 11].

In motor rehabilitation, or the recovery of motor function, MI is widely used. The categorization of MI can be based on the type of input utilized, resulting in different modalities. These include visual MI, in which the movement is visualized, and kinesthetic MI, in which the sensation of the movement is imagined, as described by Hehenberger et al. [1], Jeannerod et al. [62].

When comparing the neurophysiological aspects of kinesthetic MI and visual MI, different patterns emerge. Kinesthetic MI induces more activity in motor-associated structures and in the inferior parietal lobule, which are regions of the brain associated with motor processing. On the other hand, visual MI predominantly activates the occipital regions (associated with visual processing) and the superior parietal lobules [63, 64].

Moreover, both kinesthetic MI and visual MI have been associated with ERD and ERS [20, 65, 66, 67, 68]. These neurophysiological phenomena occur in the brain regions mentioned above, further emphasizing the specific neural correlates of kinesthetic and visual MI [64].

A study by Neuper et al. [69] compared the classification accuracies of different types of movements and found that kinesthetic MI outperformed visual MI in terms of classification accuracy. Moreover, better results were obtained when movement execution and movement observation were combined, compared to both visual and kinesthetic MI. When examining the classification patterns associated with these different tasks, both movement execution and kinesthetic MI were found to have strikingly similar patterns. Recognition of these tasks was primarily influenced by electrode positions situated over the central cortical area [1]. This suggests that neural activation patterns during kinesthetic MI and actual movement execution show similarities, especially in the central cortical region [69]. These results highlight the importance of kinesthetic MI and actual movement execution as effective strategies for motor task recognition. The similarity of classification patterns between kinesthetic MI and movement execution suggests that the underlying neural mechanisms involved in these processes may overlap to some extent. Understanding these patterns and corresponding electrode locations may aid in the development of more accurate and efficient BCI systems and rehabilitation protocols that take advantage of the shared neural networks involved in kinesthetic imagery and movement execution [1].

Given the negative effects of compromised somatosensory feedback on movement processes [39, 40], the benefits of incorporating artificial somatosensory input (either as guidance or feedback) to augment MI task performance have attracted considerable interest among researchers. Thus, the desire to create a more integrated simulation of the motor control loop has led to the exploration of different strategies to supply artificial somatosensory feedback in scenarios where MI is used as a control mechanism for end-effectors. In particular, the integration of somatosensory feedback or guidance with MI tasks has been explored in relation to ERD-derived features [1].

When comparing the classification performance of MI tasks using different feedback modalities, such as vibrotactile feedback or visual feedback, several studies have reported no significant effects of the feedback modality on the classification accuracy [49, 52]. The

vibrotactile modality was perceived as a more natural experience by the participants in research conducted by [49]. These results suggest that while somatosensory feedback may enhance the overall user experience, its effects on the classification performance of MI tasks may not be significantly different from purely visual feedback. Overall, the exploration of artificial somatosensory input in conjunction with MI promises to improve motor control and the effectiveness of MI-based applications. By accurately simulating the motor control loop and incorporating appropriate somatosensory feedback, researchers hope to bridge the gap between motor intention and execution, ultimately improving the performance and user experience of MI tasks.

Several studies in the field of rehabilitation have highlighted the beneficial effects of integrating passive movement feedback, ERD, elicited during MI, and motor scores. Specifically, Ramos-Murguialday et al. [60], Ang et al. [70] recorded enhancements in Fugl-Meyer assessment motor scores in stroke patients after MI training with feedback in the form of passive joint movements. In addition, passive movement feedback was found to improve participants' BCI performance compared with control groups that received sham feedback, as per Ramos-Murguialday et al. [59, 60]. It was also observed that performance increased when passive movement guidance was added to an MI task, as stated by Randazzo et al. [61]. In addition, Corbet et al. [56] reported increased ERD during MI with electrotactile guidance as opposed to visual guidance. Corbet et al. [56] showed that electrotactile input does not directly trigger ERD as long as it does not exceed the motor threshold. In contrast, Hommelsen et al. [71] observed similar ERD patterns in the mu frequency band when comparing a motor task with sensory-threshold electrotactile feedback and sensory-threshold electrotactile stimulation without any movement. A study by Hehenberger et al. [35] that involved actual center-out movements with real-time kinesthetic vibrotactile feedback undertook an initial investigation of the potential effects of vibrotactile input on direction decoding from EEG. However, because the focus of said study was primarily on movement decoding rather than direction decoding, the results regarding direction decoding remained largely uncertain [1].

Originally conceived in the field of thermodynamics as a metric for the disorder of a thermodynamic system, entropy was later incorporated into information theory, allowing for the quantification of the information content in a probability density function (PDF) [10, 72, 73]. When applied to nonstationary signals in the time-frequency domain, the

entropy-based estimate of signal complexity can be viewed as a measure of the concentration of the 2D energy distribution [10, 74]. The appeal of TFRs lies in their ability to offer a clear interpretation and precise measurements of actual frequencies, specific time instances of their occurrence, and a determination of whether a signal is mono-component or multi-component [10]. TFRs are categorized into Cohen's class and the Affine class Given the abundance of cross-terms in the affine class TFRs [10], the Cohen's class TFRs prove to be more appropriate for handling non-stationary multi-component signals such as EEG. In addition to the standard TFRs, reassigned versions can also be computed for many TFRs. These remapped TFRs employ the reassignment method to improve signal sharpness and concentration [75].

Numerous recent studies have explored the applications of entropy in EEG SMR for diverse objectives [33]. For instance, Zhang et al. [76] developed and applied the spectral entropy of resting-state (eyes closed) EEG as a biomarker for predicting SMR BCI performance. The use of Shannon entropy for INC state detection to reduce unintended commands during SMR BCI operations was demonstrated by Tonin et al. [77], who reported 93.70% accuracy for SMR detection prediction. Another study leveraging entropy for motion prediction and detection of the INC state was carried out by Tortora et al. [78], who reported an accuracy of 80% for motion prediction. In a study conducted by Jeong et al. [79] using the dataset from Ofner et al. [8], spectral filtering was employed to enhance detection of MRCPs, resulting in a 74% detection accuracy for "elbow flexion" movement. An impressive accuracy of 90.50% for detecting the "hand open" movement was achieved by Ieracitano et al. [80] using the same dataset. Chen et al. [81] employed entropy for feature extraction, using a combination of four entropy features (amplitude-based Shannon entropy, phase-based Shannon entropy, wavelet entropy, and sample entropy) for classifying right and left-hand MI, achieving average accuracies of up to 85.71%. Sawant et al. [82] combined empirical mode decomposition, common spatial patterns, power spectral entropy, and the Walsh-Hadamard transform to derive features, resulting in an average classification accuracy of 87.33% for right and left-hand MI. Lastly, Ji et al. [83] utilized discrete wavelet transform, empirical mode decomposition, and approximate entropy to extract features for right and left-hand MI, achieving a classification accuracy of 85.71%.

A variety of methods are used in classifying MI data, with linear discriminant analysis (LDA) and support vector machine (SVM) being regularly used. Similar results to those

of SVM and LDA were confirmed by a thorough investigation of the random forest (RF) classifier. With the advances in machine learning methods, CNNs have gained increasing attention in the BCI field owing to their potential as a classification tool. [27].

LDA is often used as a benchmark for BCI research, particularly for evaluating experimental paradigms, novel patterns, or new types of signals captured with BCIs [1, 8, 35, 56, 84]. For instance, Ofner et al. [8] reported accuracy close to chance level for a 6-class problem, while Hehenberger et al. [1] reported an average accuracy of approximately 64% when using shrinkage regularized LDA (sLDA) for classifying different MIs of the same limb.

SVM has also been shown to be very effective in classifying MI data. Vargic et al. [85] reported accuracies ranging between 47.86% and 70.71% in classifying MI from different limbs, and Ma et al. [86] reported even better results with accuracies between 76% and 91% for similar tasks of MI of different limbs.

RF classifiers have yielded comparable results. Zhang et al. [87] reported accuracies up to 76%, while Steyrl et al. [84] achieved accuracies up to 79.30% by combining RF with common spatial patterns (CSP), and Bentlemsan et al. [88] achieved 79.77% accuracy.

Recently, research with neural networks, especially for MI classification of different limbs, has increased. For example, Zhang et al. [89] achieved an impressive 90% accuracy in classifying the right and left-hand MI using a wavelet neural network (WNN) in artificially augmented data. Hou et al. [90] reported 93.06% accuracy when using graph convolutional neural networks (GCNs) to classify MI of the right hand, left hand, fists, and feet. Strahnen and Kessler [91] used a deep neural network (DNN) and achieved up to 80.7% accuracy in classifying MI of the cyclic opening and closing of the left or right fist. In contrast, Lee et al. [92] used a channel-wise variational autoencoder CNN to classify data from Ofner et al. [8] and achieved up to 60% accuracy in classifying the "elbow extension" compared to other different MIs of the same limb.

The Table 2.1 presents a comparison of MI classification accuracies across different classification methods and types of MI. The MI types are categorized into two groups: "Different-limb different MIs" and "Same-limb different MIs". From the table, it is evident that the classification accuracies for same-limb different MIs tend to be lower than those for different-limb different MIs. This discrepancy can be attributed to the more subtle variations in the signals associated with same-limb different MIs, which make it

challenging to distinguish between finer differences in movements [27].

Table 2.1: Review of MI classification accuracies by classification methods and types of MI.

	Author	Method	Accuracy (%)
Different-limb different MIs	Vargic et al. [85]	SVM	70.17
	Ma et al. [86]	SVM	91
	Zhang et al. [87]	RF	76
	Steyrl et al. [84]	RF+CSP	79.30
	Bentlemsan et al. [88]	RF	79.77
	Zhang et al. [89]	WNN	90
	Hou et al. [90]	GCN	93.06
	Strahnen and Kessler [91]	DNN	80.70
Same-limb different MIs	Ofner et al. [8]*	sLDA	23
	Hehenberger et al. [1]	sLDA	64
	Lee et al. [92]	CNN	60

*6 class problem, significance level was 18%

Chapter 3

EXPERIMENT AND DATASETS

Two different data sets were used for the research in this dissertation. The first dataset was curated as part of the original study that was conducted for this dissertation and the "Feel Your Reach" project, which was carried out by the Institute of Neural Engineering at the Graz University of Technology and was supported by the European study Council [1, 11]. In the interest of clarity and simplicity, this dataset will be referred to as the KGU dataset throughout this dissertation. This name was chosen because the data was obtained in the research where **K**inestheticc **GU**idance was incorporated into the paradigm. This particular dataset was selected due to its simplicity (two simple linear continuous center-out MIs of the same: direction *up* and direction *right*) as well as variety (two conditions: MI with visual and vibrotactile guidance and MI with visual guidance only) [27].

The second dataset is the MI dataset taken from the BNCI Horizon 2020 project [8]. For simplicity and clarity, this dataset will be referred to as the ULM dataset in this dissertation. This name was chosen because the dataset originated from the research in which **U**pper **L**imb **M**ovements were used. The ULM dataset was integrated into this research to test the developed approaches on one of the most commonly used MI datasets available online. In addition, this dataset was chosen for its simplicity (two simple linear continuous center-out MIs of the same limb were selected: *elbow flexion* and *elbow extension*). Although the ULM dataset was not originally curated as part of this dissertation, in Sections 3.1. and 3.2. the details of the data collection procedure and the experiment paradigm will be described for both datasets, as these details are essential to

the context of this dissertation.

Simple MI tasks in both datasets are relatively equivalent to one another, despite the fact that the experiment of the KGU dataset uses visual guidance and a combination of visual guidance and vibrotactile guidance, whereas the experiment of the ULM dataset uses only visual cues.

3.1. KGU Dataset

This section provides a comprehensive overview of the experimental procedures and methodologies employed in the study involving the KGU dataset [1]. It commences by detailing the participants involved, outlining their demographic information, and any selection criteria that were used. It then delves into the specifics of the experimental setup, describing the environment and equipment used, along with the process for vibrotactile stimulation. Subsequent information covers the signal acquisition process, including the types of signals recorded and the technologies employed to capture them. Lastly, the section presents the participant questionnaire, elaborating on the nature of the questions asked and the scale used for responses, offering crucial insights into the subjective experiences of the participants during the experiment.

3.1.1. Participants

The experiment was performed with 15 able-bodied participants (seven male, eight female; age 21-32). All participants were self-reportedly right-handed. Participants for the experimental study were recruited using a convenience sampling methodology. The pool of participants was sourced from several distinct groups: acquaintances and colleagues of the authors of the study, members of the laboratory staff at the Institute of Neural Engineering - Graz University of Technology, students enrolled in the Institute of Neural Engineering, and respondents to promotional posters disseminated across the Institute of Neural Engineering. This approach allowed for a diverse range of participants while also maintaining the necessary control over the experimental conditions.

Ten participants self-reported prior experience with MI, and six participants self-reported prior experience with MI and vibrotactile stimulation. Participants received

both written and verbal instructions prior to the experiment and provided written informed consent for participation. Participants were compensated 7.50 euros per hour for their participation. The study protocol was approved by the Ethics Committee of the Medical University of Graz [1].

3.1.2. Experimental Setup

Participants were seated in front of a computer screen (a particular BCI setup can be seen in Fig. 1.3) and were instructed to perform a guided MI of a center-out arm movement task. The movement they imagined performing was moving their right palm slowly across the flat surface of a desk. The imagery was guided either by a visual moving cue displayed on the computer screen and a simultaneous vibrotactile moving sensation across the right shoulder blade (condition VtG: MI with vibrotactile guidance) or by a visual moving cue alone (condition noVtG: MI without vibrotactile guidance). The direction of the center-out movement was given in one of two orthogonal directions: to the right or up/forward. Participants imagined that they moved their palm across the table to the right or to the front/forward. In the second case, the visual and vibrotactile guidance moved upward, while the imagined movement was to the front/forward. In the following, this will be referred to as the *up* direction [1].

At the beginning of the experiment, a familiarization procedure was conducted to ensure that all participants, including those without prior experience with vibrotactile stimulation or MI, were adequately prepared. This approach aligns with the recommendation proposed by Roc et al. [93] to effectively familiarize users with MI tasks. The familiarization process involved conducting practice runs for each condition and providing participants with the opportunity to practice both executing the center-out movement on a table surface and imagining the same movement. This approach aimed to facilitate participants' vivid memorization of the movement pattern [1].

After the familiarization phase, each participant completed a total of three runs per condition in a block design (three consecutive runs for each condition). The order of conditions was counterbalanced among participants, with eight participants starting with the VtG condition and seven participants starting with the noVtG condition. Each run consisted of 40 regular trials, except for the VtG condition runs, which included an addi-

tional four trials where the movement direction indicated by the vibrotactile guidance and the visual guidance were incongruent. In the VtG condition, participants were prompted to provide a rating after each trial to indicate whether the two guidance modalities were congruent or not (whether they moved in the same direction). The incongruent trials were primarily used to maintain participant engagement and were not included in the subsequent classification and analysis stages. Each trial had a duration of 7.5 s, with the MI task being performed for a period of 2 s, as depicted in Fig. 3.1. The visual cues, represented in the middle row of the figure, were identical for both conditions. The top row is highlighted in green and blue to denote the VtG and noVtG conditions, respectively. The circles depicted in the sketch on the right indicate the tactors used, with active tactors marked in purple and idle tactors marked in white. The purple arrows represent movement in both visual and vibrotactile representations, with an example illustrating movement to the right. To signal the start of each trial, a visual alert was presented on the screen 1.5 s prior to the appearance of the fixation cross. The fixation cross remained on the screen for 2 s, with the final 1.5 s of this period serving as the baseline (rest state) period. During this period, participants were instructed to fixate their gaze on the cross and relax. Subsequently, the visual cue, depicting a right hand with a fixation point (a black dot in the middle of the hand), appeared on the computer screen. In the VtG condition, vibrotactile stimulation was engaged simultaneously with the visual cue's appearance. The cue and the vibrotactile stimulation remained stationary for a pre-MI period of 2 s, after which they moved either to the right or upwards at a constant speed. Participants were instructed to perform the MI task in accordance with the movement direction indicated by the cue and the vibrotactile stimulation. In the VtG condition, participants were then prompted to assess whether the vibrotactile guidance provided in that trial was congruent with the visual guidance. They were required to respond by pressing a designated key. The question appeared on the screen 1 s after the completion of the task and remained visible until the participant provided a response [1].

Following the completion of the main runs, two additional runs of continuous rest were recorded, each lasting for one minute. Subsequently, participants underwent two runs specifically designed to elicit controlled eye artifacts, including saccades and blinks, as described by Kobler et al. [94]. Both rest and artifact runs were later used in the preprocessing of the data.

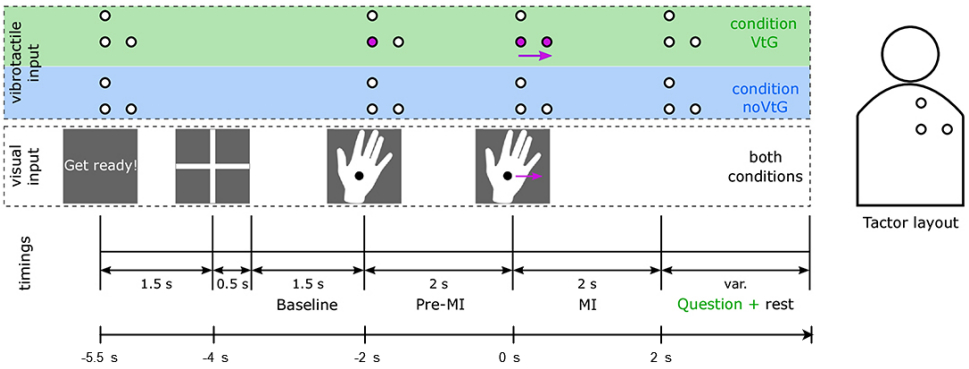


Figure 3.1: Trial structure and tactor layout according to the experiment paradigm of the KGU dataset [1].

The experiment was set up using a variety of specialized tools, each playing a key role in the design and execution of the experiment.

The Simulation and Neuroscience Application Platform (SNAP) was the primary software suite employed in implementing the experimental paradigm. SNAP is an open-source, high-level software platform, hosted on Github [95], that offers tools for designing, conducting, and analyzing neuroscience experiments. As a flexible and comprehensive tool, SNAP is widely used in the neuroscience community to conduct experiments across a broad range of research topics.

SNAP operates on the Python 2.7 programming language, which is well-regarded for its readability, simplicity, and wide range of supported libraries. Python 2.7 was specifically chosen for its extensive support and its robust collection of scientific computation packages, which facilitate the management of large and complex datasets.

Lastly, the Panda3D game engine [96] played a crucial role in the visual aspects of the experiment. Panda3D, a free and open-source platform, enables the development of 3D simulations and interactive applications. The software, developed by Disney and Carnegie Mellon University's Entertainment Technology Center, is known for its powerful rendering capabilities and its intuitive Python interface, which aligns well with the SNAP platform. In the context of this experiment, Panda3D was employed to create immersive visual stimuli that participants interacted with during the MI tasks.

3.1.3. Vibrotactile Stimulation

To stimulate the right shoulder blade, three specialized tactile actuators, specifically C-2 tactors from Engineering Acoustics Inc. (Casselberry, USA), shown in Fig. 3.2, were affixed to the inside of an elastic shirt. These tactors were connected to a custom device equipped with an ARM Cortex M4 micro-controller from STMicroelectronics (Geneva, Switzerland). The signal driving the tactors operated at a carrier frequency of 250 Hz. The control of the tactors involved manipulating the intensities of the C-2 tactors via a serial interface using Python 2.7 and SNAP [1].

Figure 3.2: Four C-2 tactors attached to the inside of an elastic shirt and connected to an ARM Cortex M4 micro-controller.

Given the range of sensitivity profiles among individuals, the tactor amplitudes were

carefully calibrated to ensure uniform perceived intensities across all participants. This calibration took into account the varying degrees of sensitivity in different individuals, adjusting the amplitude of vibrations to achieve a consistent and equal level of perceived intensity, irrespective of innate tactile sensitivity. This step was crucial to maintaining the consistency of the delivered stimuli across the participants and ensuring the uniformity of the experimental conditions. The vibrotactile guidance in this experiment was delivered through a moving virtual stimulus designed to mimic the trajectory of the imagined limb movement. This virtual stimulus "moving" stimulus traveled from the central tactor to one of the outer tactors, following the path that the corresponding limb would take during an actual movement and enhancing the sense of kinesthesia, encouraging the brain to more accurately imagine the specific movement being examined. The tactor amplitudes were modulated to elicit a sensation of movement using the following mapping [35, 97, 98, 99]:

$$x_v = \frac{A_2^2}{A_1^2 + A_2^2} \tag{3.1}$$

$$A_v^2 = A_1^2 + A_2^2 \tag{3.2}$$

where x_v is the location of the moving (virtual) stimulus between the start tactor $T_1(x_1 = 0)$ and the end tactor $T_2(x_2 = 1)$, A_v is the amplitude of the moving stimulus, and A_1, A_2 the amplitudes of T_1 and T_2, respectively [1]. From the above equations, the amplitudes of A_1 and A_2 can be calculated based on the desired x_v and A_v:

$$A_1^2 = A_v^2 - x_v A_v^2 \tag{3.3}$$

$$A_2^2 = x_v A_v^2 \tag{3.4}$$

3.1.4. Signal Acquisition

EEG and EOG were recorded from 64 actiCap active electrodes using two BrainAmp amplifiers (Brain Products GmbH, Gilching, Germany), at a sampling rate of 1 kHz. Electrodes were arranged according to the international 10/20 EEG system [100] shown in Fig. 3.3, where 61 channels were used for EEG and 3 channels were used for elec-

trooculogram (EOG) [1].

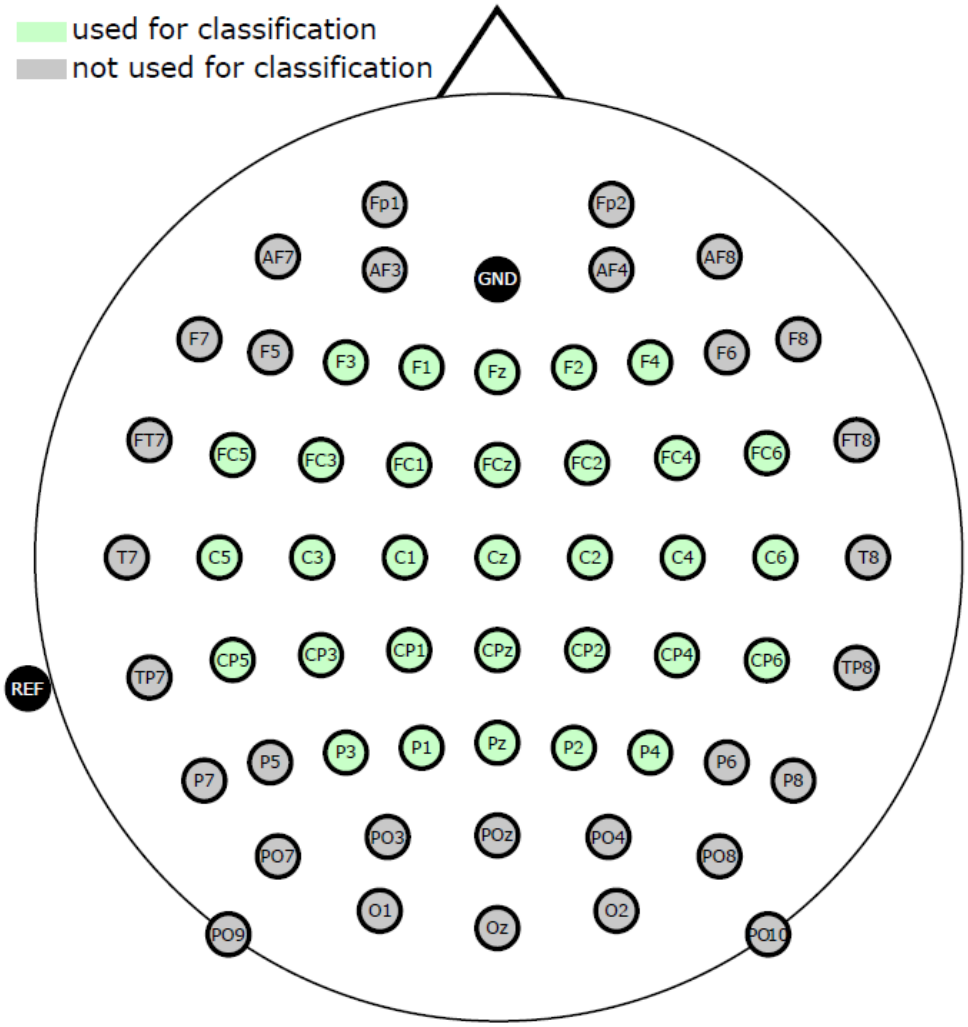


Figure 3.3: International 10/20 EEG system cap montage [1].

3.1.5. Participant Questionnaire

Upon completion of the experiment, participants were requested to complete a questionnaire. The questionnaire comprised of 7 questions for the VtG condition and 5 questions for the noVtG condition. Participants provided their responses by selecting values on a Likert scale ranging from 1 to 5. The scale answers were: "Strongly disagree" (1), "Disagree" (2), "Neither agree nor disagree" (3), "Agree" (4), and "Strongly agree" (5). The questionnaire can be found in the dissertation Appendix A..

3.2. ULM Dataset

In this section, the ULM dataset [8] is described. As mentioned in the beginning of Chapter 3, this dataset was not curated as part of this dissertation, but it is still important to describe the experiment paradigm and recording to a certain extent, as they have an impact on the differences between the results acquired by applying the same methods to different datasets. For this reason, the experimental setup of the ULM dataset will not be explained in great detail like it was for the KGU dataset. The section starts by presenting the participants involved and proceeds to give a concise description of the experiment. Next, the essence of the experiment and its paradigm are briefly outlined. The section then concludes with a description of the signal acquisition process, detailing the types of signals recorded and the technologies used for data capture.

3.2.1. Experimental Description

Data were collected from a total of 15 participants, including 6 men and 9 women, with ages ranging from 22 to 40 years (mean age = 27, standard deviation = 5). The majority of participants (14 out of 15) were right-handed. During the MI session, each participant completed ten runs, with each run consisting of six different MI tasks. The MI tasks included elbow flexion, elbow extension, supination, pronation, hand closing, and hand opening. Static visual cues corresponding to these tasks were presented on the screen in front of the participant. Each run comprised a total of 36 trials, with six trials for each of the six tasks. As depicted in Fig. 3.4, each trial had a duration of 5 s, and the MI period occurred within a 3 s interval. The fixation cross was displayed for 2 s, with the last 1.5 s later utilized as the baseline for signal processing and classification within this research. Participants were instructed to maintain their gaze fixed on the fixation cross during this interval and to minimize movements in order to reduce artifacts. Subsequently, a stationary visual cue indicating one of the six movements was presented on the computer screen, which instructed participants to perform the MI task based on the given cue [8]. In the present research, only two of the aforementioned tasks, namely *elbow flexion* (EF) and *elbow extension* (EE), were utilized. These tasks were selected due to their similarity to the tasks employed in the KGU dataset, specifically the MI tasks of directions *up* and

right.

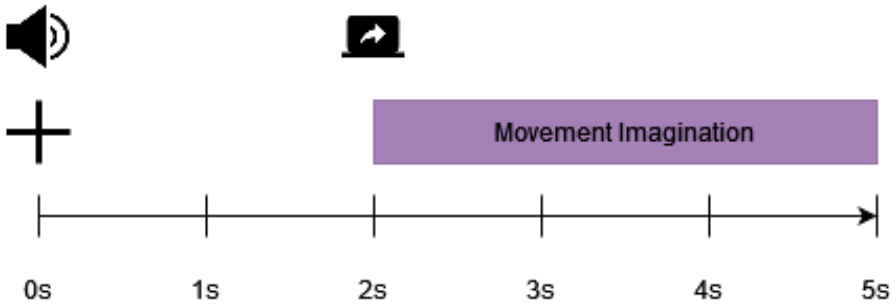


Figure 3.4: Experiment paradigm of the ULM dataset [33, 8].

3.2.2. Signal Acquisition

For data acquisition, four g.tec amplifiers from g.tec medical engineering GmbH (Austria) were utilized. The EEG signals were recorded from 61 active electrodes, while the EOG signals were captured from 3 active electrodes. The amplifiers operated at a sampling rate of 512 Hz. To ensure the appropriate frequency range for analysis, an 8th-order Chebyshev bandpass filter was applied to the dataset, allowing frequencies between 0.01 Hz and 200 Hz to pass through. Among the electrodes, only the 31 electrodes highlighted in green (as shown in Fig. 3.3) conformed to the international 10/20 EEG system cap montage and were in congruence with the KGU dataset montage described in Section 3.1.. Consequently, for signal processing, analysis, and classification purposes, only the corresponding EEG and EOG electrodes were used for the ULM dataset.

Chapter 4

SIGNAL PROCESSING

In the context of EEG, signal processing plays a crucial role in deciphering the complex patterns and information contained in the recorded brain waves. EEG signals are often contaminated with artifacts and noise, stemming from various sources such as muscle activity, eye movements, or power-line interference. Thus, signal processing techniques are required to enhance the signal quality and extract meaningful features from this data. These techniques include filtering to remove the unwanted frequencies, artifact removal to reduce the impact of non-brain signal sources, and feature extraction to transform the EEG data into a form that can reveal underlying neural processes or be utilized for classification in BCIs. The ultimate goal of EEG signal processing is to enhance understanding of brain activity and facilitate its practical applications, such as in neurofeedback, cognitive neuroscience, and BCIs [5]. As part of this dissertation, various types of processed features were analyzed. Namely, low-frequency amplitude features around the delta range ($0.2 - 5\ Hz$), broad-frequency amplitude features ($1 - 40\ Hz$), spectral power features with a focus on the mu ($8 - 12\ Hz$) and beta ($15 - 32\ Hz$) frequency bands, and low-frequency short-term Rényi and Shannon entropy features based on different TFRs. The processing of all said features is described in the following sections.

4.1. Preprocessing and Amplitude Features

Prior to the classification phase, both datasets were preprocessed with the following steps [1]:

1. The data were downsampled to 200 Hz, a 4th-order zero-phase Butterworth band-pass filter was utilized with a passband between 1 and 40 Hz, to remove noise and unwanted frequency components, focusing on the relevant frequency range for MI.

2. Data were epoched to the relevant time period: from $t = -5.5$ s to $t = 2$ s as shown in Fig. 3.1 for the KGU dataset and from $t = 0$ s to $t = 5$ s as shown in Fig. 3.4 for the ULM dataset [1, 27].

3. The EEGLAB Matlab toolbox was employed to exclude trials that demonstrated artifact presence and exceeded the amplitude threshold [7]. For the identification of artifact-contaminated trials, individual raw signals were subjected to band-pass filtering between 1 and 60 Hz. These were then scrutinized based on maximal amplitude, kurtosis, and joint probability. Trials that displayed amplitudes exceeding 200 μV, or those where either kurtosis or joint probability exceeded five times the standard deviation, were marked for rejection [1].

4. Utilizing the EEGLAB toolbox, independent component analysis (ICA) [7, 101] was carried out individually for each participant. For the KGU dataset, ICA was conducted on 61 EEG channels, resulting in 61 independent components, with 3 EOG channels used for the removal of artifacts. For the ULM dataset, the analysis was done on 31 EEG channels (corresponding to the same positions as the KGU dataset), generating 31 independent components, with 3 EOG channels reserved for artifact elimination. For both datasets, only independent components (IC) related to MI were retained using the Infomax algorithm [101, 102] in combination with SASICA [103] and manual IC rejection by visual inspection. The ICA weights were saved and subsequently applied to the data [1].

5. Data were further filtered and downsampled in the bands of interest using a 4th-order zero-phase Butterworth bandpass filter, specifically with a passband between 0.2 and 5 Hz for low-frequency features (downsampled to 20 Hz), and a passband between 1 and 40 Hz for broad-frequency features (downsampled to 100 Hz) [27]. Low-frequency and broad-frequency features are subsequently used separately for different classification and processing purposes.

6. For both datasets, 31 matching relevant channels around the sensorimotor cortex were used for further preprocessing and analysis (marked in green in Fig. 3.3).

An overview of the common preprocessing (applied to both datasets) is shown in Fig. 4.1. The low-frequency and broad-frequency preprocessed features described above were utilized as amplitude features for classification. Subsequently, additional processing methods were applied depending on the purpose of further analysis, in particular for spectral power features and information entropy features, both of which are described in the following sections.

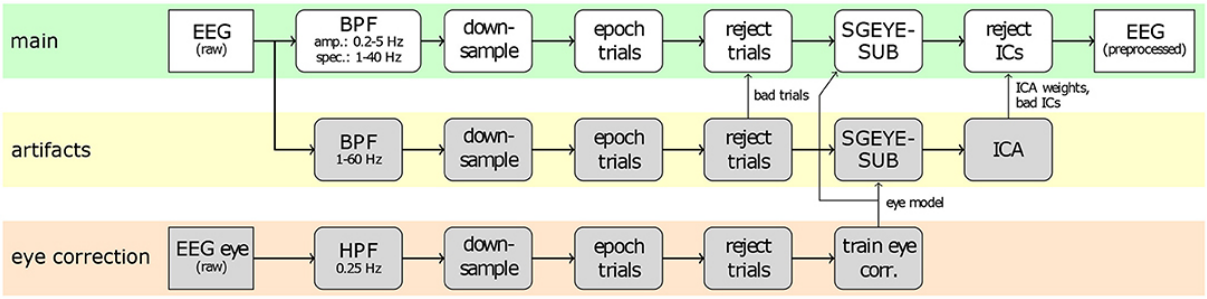


Figure 4.1: Common preprocessing techniques that were applied to EEG signals in this research. The uppermost branch, colored green, depicts the primary preprocessing techniques that were implemented on the trials. The lower branches illustrate the methods utilized to remove eye artifacts (colored in orange at the bottom) and general artifacts (colored in yellow in the middle) [1].

4.2. Spectral Power Features

For the initial research segment, the *KGU research*, in addition to amplitude features, spectral power features were investigated [1]. The broad-frequency features acquired from the preprocessing phase were subjected to further filtering with a 4th-order zero-phase Butterworth bandpass filter, specifically in the mu ($8 - 12\ Hz$) and beta ($15 - 32\ Hz$) frequency bands. To amplify the class separability, CSP were computed for each participant [104, 105, 106]. These CSP filters are designed to enhance the variance of spatially projected power signals for one set of features while minimizing it for the opposite set of features. Sets of features represent classes of different MI directions (*up*, *right*), different conditions (VtG, noVtG), or MI and rest (baseline). The calculation of these filters was carried out within a time window during the MI period, i.e., 0.5 to

1.5 s following the cue motion onset. Once the top five most distinguishing CSP filters for each of the two sets of features were applied to the data (ten features in total), the logarithmic power relative to the baseline period of each trial was calculated for each feature separately within the mu and beta frequency bands. Both frequency ranges were captured by combining these features. Finally, a moving-average filter with a window size of 1 s was applied to the combined features to derive the classification features. This is a smoothing process aimed at noise reduction and enhancement of the stability of the classification features [1].

4.3. Signal Processing for Neurophysiology Analysis

To provide a more comprehensive understanding of the neurophysiological aspects underlying the observed low-frequency amplitude and spectral power features, additional analyses were conducted [1]. These analyses aimed to investigate the spatial and temporal characteristics of neural activity during MI. The following two steps were further performed:

1. In the case of low-frequency ($0.2 - 5$ Hz) amplitude features, the preprocessed signals underwent re-referencing to the common-average-reference. This step helps to reduce the influence of common noise sources and emphasizes the specific neural activity associated with MI [1], which is very practical when representing the data visually.

2. For spectral power features, a time-frequency decomposition was performed using Morlet wavelets [107]. This technique allows for a detailed examination of the signal's spectral content over time. The Morlet wavelets were applied with a full width at half maximum (FWHM) of 3 s at 1 Hz resolution, covering frequencies from 1 to 40 Hz. By analyzing the signal in 1 Hz steps, a comprehensive representation of frequency-specific modulations throughout the entire frequency range of interest was obtained.

4.4. Information Entropy Features

Entropy research segment done as part of this dissertation focused on the utilization of short-term information entropy for effective MI detection [33]. To achieve this, a series of TFRs were computed from amplitude features. Subsequently, short-term entropy measures, namely Rényi and Shannon entropy, were computed from the aforementioned TFRs represented as PDFs. Two distinct window sizes were employed in the process of calculating such entropy features, namely a long window ($w = 1\ s$) and a short window ($w = 0.5\ s$), which was done to investigate the impact of varying window lengths on the resulting entropy measures. The moving step for each window size was one sample. The research done as part of this segment involved a comparative analysis of amplitude features, Shannon entropy, and Rényi entropy features for classification purposes, as described in Chapter 5.

4.4.1. Information Entropy

The concept of entropy, originally derived from the field of thermodynamics as a measure of the disorder or randomness within a thermodynamic system, has found broader applications beyond its initial domain. In particular, the introduction of entropy to the field of information theory has allowed for the quantification of the information content or uncertainty present in a probability distribution or a dataset [33, 10]. This extension of entropy has enabled researchers to explore and analyze the complexity, organization, and patterns within various types of data, ranging from physical systems to information signals and beyond. By harnessing the power of entropy-based measures, researchers have gained valuable insights into the underlying structures, dynamics, and information content of complex systems, paving the way for advancements in diverse fields such as signal processing, data analysis, machine learning, and neuroscience [10]. At its core, information entropy is a statistical measure that quantifies the uncertainty, randomness, or unpredictability of a dataset. The higher the entropy, the greater the uncertainty or randomness in the data. Information entropy can be used to describe the complexity of information content in data signals, such as EEG signals. For example, in the context of BCIs, information entropy can potentially provide valuable insights into brain activ-

ity patterns and be used to enhance the performance of classification algorithms. This tool transforms complex, high-dimensional data into simpler, more interpretable metrics, thereby aiding in the process of deciphering intricate neural signals. Information entropy has enabled the quantification of the information content represented in a PDF [10, 33, 72, 73]. The estimation of signal complexity in nonstationary signals on a time-frequency plain can be seen as a concentration of energy distribution in a 2D space, and as such, it is used by entropy-based methods [10, 74], effectively meaning that TFRs can be interpreted as PDFs.

This segment of the dissertation presents a comparison of amplitude and entropy features, derived from various TFRs, in EEG MI datasets where visual guidance alone, a mix of visual and vibrotactile guidance, or solely visual cues accompanied the execution of MI by participants. After undergoing signal processing and feature extraction, these features were evaluated based on their classification accuracy and F1 score performance upon classification. The primary driving force behind this research was to assess the utility of short-term information entropy, obtained from a variety of TFRs, for more effective detection of MI, hence enhancing INC state detection. This could potentially augment the overall performance of MI-controlled BCI systems. Further, the role of vibrotactile guidance in MI detection was investigated, providing valuable insights into possible enhancements for BCI systems that incorporate somatosensory input [33].

The information entropy is used as an indicator of the energy distribution concentration of the TFR [72]. Its application allows assessment of the level of concentration or dispersion of signal components present in the TFR. In the context of TFR analysis, a highly concentrated TFR with a small number of signal components exhibits lower entropy compared to a TFR that contains a larger number of signal components [108]. This relationship provides a means to characterize the level of information content and structural organization within the TFR.

To compute the entropy, a short-term approach was employed, utilizing the moving window technique. The width of the window was determined based on two options: a long window of 1 s or a short window of 0.5 s. A window size 0.5 s was chosen as the lower limit due to the marginal improvement in accuracy with further reductions in window size, which is outweighed by the increase in computational time [33]. The window was moved in 50 ms steps across each trial and computed separately for each channel. This approach

allowed for the evaluation of entropy at different time points, taking into account the equidistant segments on both sides of the targeted time-point (25 ms on each side) [33].

Shannon entropy is defined as the negative sum (or integral for a continuous stream of variables such as TFR) of the probabilities of each possible outcome of the variable multiplied by the logarithm of that probability [10], defined as:

$$H_x = -\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} TFR_x(t,f)\log_2 TFR_x(t,f)dtdf \tag{4.1}$$

where TFR represents a PDF of t (time) and f (frequency).

To compare with the Shannon entropy, Rényi was also utilized. Rényi entropy is a generalization of Shannon entropy. It introduces a parameter α that allows for different degrees of emphasis on rare or common events in the distribution. The Rényi entropy is calculated as:

$$R_x^\alpha = \frac{1}{1-\alpha}\log_2\left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} TFR_x^\alpha(t,f)dtdf\right) \tag{4.2}$$

where R_x^α is the α order Rényi entropy for the TFR_x (due to oscillation-reducing effects, Rényi entropy order was set to $\alpha = 3$ as in [73]). Rényi entropy includes Shannon entropy as a special case when α approaches 1. When α takes on other values, Rényi entropy provides different perspectives on the distribution. For example, for small values of α, Rényi entropy emphasizes rare events and is more sensitive to outliers in the distribution, while for large values of α, it places more emphasis on common events. Since certain TFRs can assume negative values due to interferences (or even magnitude and odd phase) [10], absolute values of the calculated TFR before calculating an entropy were taken [33]. Examples of both Shannon and Rényi entropy results applied on sample TFRs can be seen in Section 6.2..

4.4.2. Time-frequency Representations

TFRs offer the ability to interpret and accurately measure real-time frequencies and their occurrence time-points, additionally demonstrating whether the signal is mono-component or multi-component [10]. Despite the availability of a range of TFRs for diverse applications, this research focuses on those TFRs that are most suitable for characterizing the employed datasets and calculating Rényi and Shannon entropy [33].

Two broad categories classify TFRs: Cohen's class and Affine class. TFRs in Cohen's class are quadratic or bilinear, demonstrating covariant properties under time and frequency translation. Conversely, the Affine class TFRs are bilinear and reveal covariant characteristics by translating in time and dilating in frequency. However, due to the excessive presence of cross-terms in the Affine class TFRs [10], this research prioritizes Cohen's class TFRs. Several TFRs from Cohen's class and their corresponding reassigned versions were scrutinized. Reassigned TFRs apply the reassignment method to enhance signal sharpness and concentration. The objective of the reassignment method is to relocate TFR's values towards the center of gravity, thereby achieving superior signal component localization [75]. This method's fundamental concept is that a specific distribution's values may not necessarily be symmetrically distributed around a certain time-frequency point, where they are traditionally calculated. Instead, they tend to be located at this domain's center of gravity, providing a more accurate depiction of the signal's local energy distribution [109]. In the context of this research, six distinct TFRs and their corresponding reassigned versions were interpreted as two-dimensional PDFs. They were used as inputs for Rényi and Shannon entropy, essentially facilitating the analysis of TFRs' complexity and information content [33].

For the entropy calculation, following Cohen's class TFRs were utilized, all implemented through Matlab *Time-frequency toolbox* [109, 33]:

1. Spectrogram, which is a simple Cohen's class TFR and can be interpreted as a bilinear energy distribution. Spectrogram has a trade-off between time resolution and frequency resolution as a drawback but good interference deduction if two signal components are sufficiently far apart [33, 110, 109]. When the spectrogram is computed with a shorter time window, it offers better time resolution, allowing for the detection of rapid changes in the signal. On the other hand, using a longer time window improves frequency resolution, enabling the identification of narrowband frequency components. One of the drawbacks of the spectrogram is that it may not accurately represent signals with rapidly changing frequency content or non-stationary characteristics. In such cases, the spectrogram may exhibit blurring or smearing of frequency components over time, making it challenging to precisely

identify individual frequency components. The spectrogram is calculated as:

$$S_x(t,\nu) = \left| \int_{-\infty}^{+\infty} x(u) h^*(u-t) e^{-j2\pi\nu u} du \right|^2 \tag{4.3}$$

where h is a frequency smoothing window. The spectrogram can be interpreted as a measure of the energy of the signal contained in the time-frequency domain centered on the point (t, ν) [109].

2. Reassigned spectrogram, introduced as an attempt to improve the spectrogram's localization to produce a sharper representation of signal components [109]. The reassigned spectrogram is calculated with the equation [33]:

$$\begin{aligned} RS_x\left(t', \nu'; h\right) = \int \int_{-\infty}^{+\infty} S_x(t, \nu; h) \delta\left(t' - \hat{t}(x; t, \nu)\right) \\ \delta\left(\nu' - \hat{\nu}(x; t, \nu)\right) dt d\nu \end{aligned} \tag{4.4}$$

where δ is the reassignment operation, (t', ν') is the value of the reassigned spectrogram, and $(\hat{t}, \hat{v})$ is the center of gravity of the signal energy distribution around (t, v). The reassigned spectrogram also uses the phase information of the short-time Fourier transform and not only its squared modulus, as is the case with the spectrogram [109].

3. Gabor representation was introduced to remove the highly oscillated cross-terms without significantly altering desirable properties, i.e., it can balance the resolution and cross-term interference [10, 33, 111]. These cross-terms can arise when multiple signal components overlap in the time-frequency domain, leading to interference and difficulties in accurately distinguishing between individual components. The Gabor representation combines the advantages of both time and frequency localization by using a window function known as the Gabor function or Gabor wavelet. The Gabor representation is calculated as [33]:

$$G_x[n, m; h] = \sum_k x[k] h^*[k-n] \exp[-j2\pi mk] \tag{4.5}$$

where $G_x[n, m; h]$ are Gabor coefficients (n, m). Each coefficient contains infor-

mation relative to the time-frequency content of the signal around the observed time-frequency location.

4. Reassigned Gabor spectrogram, which is a reassigned spectrogram utilizing Gabor representation [111]. Calculated with the Eq. (4.4), but utilizing a Gaussian window instead of a frequency smoothing window, thus allowing faster computation [33, 109].

5. Pseudo Wigner-Ville distribution is based on Wigner-Ville distribution (WVD). The WVD is known for its desirable properties, including the preservation of time and frequency shifts as well as energy conservation. However, one drawback of the WVD is that it often produces strong cross-terms in multicomponent signals [10, 33], which can hinder the accurate representation and analysis of individual signal components. To overcome this limitation, the Pseudo Wigner-Ville distribution introduces a windowing operation to the WVD [109, 112]. This windowing operation is equivalent to applying frequency smoothing to the WVD and serves to attenuate the cross-terms that arise in multicomponent signals. By incorporating the windowing operation, the Pseudo Wigner-Ville distribution is able to reduce the interference between different frequency components, allowing for a clearer representation of individual components in the time-frequency domain. It is calculated as:

$$PW_x(t,\nu) = \int_{-\infty}^{+\infty} h(\tau)x(t+\tau/2)x^*(t-\tau/2)e^{-j2\pi\nu\tau}d\tau \tag{4.6}$$

where h is the frequency smoothing operation.

6. Smoothed pseudo Wigner-Ville is a pseudo Wigner-Ville distribution that utilizes time and frequency smoothing (in contrast to frequency only smoothing that is present in Pseudo Wigner-Ville) in order to smooth the signal in the time and frequency domains [113]. The previous compromise of the spectrogram between time and frequency resolutions is now replaced by a compromise between the joint time-frequency resolution and the level of cross-terms (more smoothing results in poorer resolution) [33, 109]. It is defined as:

$$\begin{aligned} SPW_x(t,\nu) = \int_{-\infty}^{+\infty} h(\tau) \int_{-\infty}^{+\infty} g(s-t)x(s+\tau/2) \\ x^*(s-\tau/2)ds\, e^{-j2\pi\nu\tau}d\tau \end{aligned} \tag{4.7}$$

where g is the time smoothing operation.

7. Reassigned pseudo Wigner-Ville is a Pseudo Wigner-Ville TFR that utilizes the reassignment method [109], and is calculated as:

$$RPWV_x\left(t',\nu';h\right)=\iint_{-\infty}^{+\infty}PWV_x(t,\nu;h)\\ \delta\left(t'-\hat{t}(x;t,\nu)\right)\delta\left(\nu'-\hat{\nu}(x;t,\nu)\right)dtd\nu \tag{4.8}$$

8. Reassigned smoothed pseudo Wigner-Ville is a Pseudo Wigner-Ville TFR that utilizes the reassignment method and a separable (time and frequency) smoothing function. It is defined as:

$$RSPWV_x\left(t',\nu';g,h\right)=\iint_{-\infty}^{+\infty}SPWV_x(t,\nu;g,h)\\ \delta\left(t'-\hat{t}(x;t,\nu)\right)\delta\left(\nu'-\hat{\nu}(x;t,\nu)\right)dtd\nu \tag{4.9}$$

where g is the time smoothing window.

This chapter concludes with a comprehensive flowchart, depicted in Fig. 4.2, providing a visual representation of the described methods [33]. This includes the stages of data acquisition, preprocessing, classifier training, and classification, each of which is color-coded for clarity. The data acquisition phase is presented in purple, the preprocessing phase in yellow, amplitude feature processing in blue, entropy feature processing in red, power spectral power feature processing in magenta, and the training and classification phase (which is described in the following chapter) in gray and brown.

It is important to note the distinction between the data acquisition of the two datasets used in this research. The KGU dataset, annotated with green text in the flowchart, utilizes both visual guidance and a combination of visual guidance and vibrotactile guidance. Conversely, the ULM dataset, marked in orange, relies on visual cues only. A clear overview of the processing pipeline described in this chapter is provided through this graphic illustration, setting the stage for the discussions in subsequent chapters.

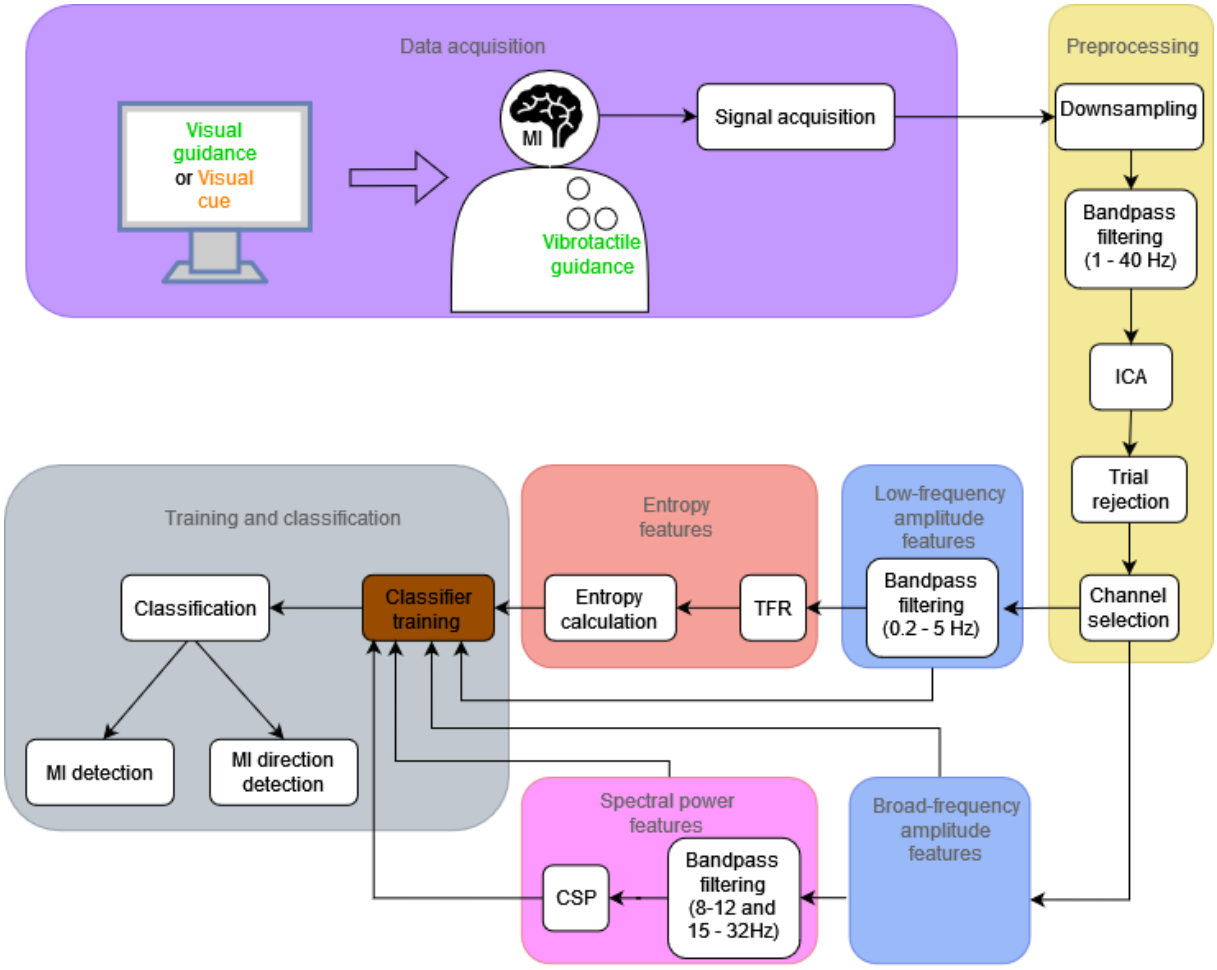


Figure 4.2: Flowchart of the processing pipeline [33].

Chapter 5

CLASSIFICATION

The process of classification plays a fundamental role in the analysis and interpretation of data across a wide array of scientific disciplines. In the context of this dissertation, classification is employed as a powerful tool to make sense of complex patterns within tested datasets, allowing for meaningful conclusions and insights to be drawn. This particular approach involves the categorization of data into predefined groups or classes based on shared characteristics or attributes. With advances in machine learning and computational power, classification techniques have become increasingly sophisticated and capable of handling large datasets consisting of multi-dimensional signals. The following section will delve deeper into the specific methods of classification utilized in this dissertation, elaborating on their principles, applications, and implications.

In the research done as part of this dissertation, classification was done for each separate research segment. First, classification of low-frequency amplitude features and broad-frequency spectral power features, described in Section 5.1., for the *KGU research* segment was done. Second, classification of entropy features, described in Section 5.2., for *Entropy research* segment was done. The third and final classification of low-frequency amplitude and broad-frequency amplitude features for the purposes of comparison of various classification methods in the *Classification methods comparison research* segment, described in Section 5.3. was done. The results of all three segments are described in Chapter 6.

The classification process in this study was performed using a 5-fold cross-validation approach. Cross-validation is a widely used technique in machine learning and pattern recognition to assess the performance and generalization ability of a classifier. In a 5-fold

cross-validation, the dataset is divided into five equal-sized subsets, or folds. The model is then trained and evaluated five times, each time using a different fold as the testing set and the remaining four folds as the training set. In instances of training and testing, 80% of the dataset was used for training the classifier and performing cross-validation. This training set allowed the classifier to learn the underlying patterns and relationships in the data. The remaining 20% of the dataset, referred to as the testing set, was reserved for evaluating the performance of the trained classifier. Each fold served as the testing set exactly once. This ensured that every sample in the dataset was used for testing once and was also part of the training set four times. By averaging the performance results from the five folds, a more reliable estimate of the classifier's performance for the unseen data was obtained.

The calculation of accuracies and F1 scores was conducted in a three-step process: Firstly, the average accuracy, or F1 score, for each participant was computed individually. Secondly, a grand-average (GA) accuracy or F1 score across all participants was deduced (this constitutes the average accuracy or F1 of all participants). Lastly, grand-average accuracy, or F1 score was considered the ultimate result of classification accuracy, or F1 [33]. The accuracy represents the percentage of correctly classified samples, while the F1 score combines precision and recall to provide a balanced measure of classification performance.

The accuracy is calculated as:

$$acc = \frac{TP + TN}{n} \tag{5.1}$$

and F1 is calculated as:

$$F1 = \frac{TP}{TP + 0.5(FP + FN)} \tag{5.2}$$

where n is the number of trials, TP is the number of True Positives, TN is the number of True Negatives, FP is the number of False Positives, and FN is the number of False Negatives. The F1 score represents the harmonic mean of the precision and recall of the classification [33].

For all classification methods and features tested, classification was performed using the same features for each subject separately. The features were obtained from the signal processing phase (Section 4.1.). Each of the two MI classes had 120 trials for the KGU

dataset (60 for the VtG condition and 60 for the noVtG condition) or 60 trials for the ULM dataset [27].

5.1. Classification in KGU Research Segment

In the initial segment, the *KGU research*, classification was performed across three distinct aspects [1]. The first aspect involved the classification of two directions, namely *up* and *right*, separately for each condition (VtG and noVtG). By comparing the classification results between the two conditions, it was possible to assess the effectiveness of vibrotactile guidance in enhancing the ability to discriminate between these specific directions.

The second aspect focused on the classification of the MI period against the baseline (rest state), independently for each condition. This analysis aimed to investigate the influence of vibrotactile guidance on the detection of the MI state compared to the resting state. The classification was performed within fixed time periods: from 0.5 to 1.5 s after the cue movement onset for the MI period and from -3.5 to -2 s before the cue movement onset for the baseline. These time periods were chosen to capture the relevant neural activity associated with MI and the resting state, as illustrated in Fig. 3.1.

Finally, the third aspect involved the classification of the two conditions (VtG and noVtG) against each other, considering the entire trial duration. This classification aimed to provide a comprehensive comparison between the two conditions throughout the entire experimental trial. By examining the classification performance between the VtG and noVtG conditions, the research sought to determine the detectability of vibrotactile guidance in such MI task.

The sLDA classifier [114], which is detailed further in Section 5.3., was utilized for all classifications in this segment of the research. This classifier was utilized because it is commonly applied in similar research [8, 22, 35, 114]. Classifications were conducted once utilizing low-frequency amplitude features and once employing spectral power features, both of which were described in Section 4.1..

5.1.1. Feature Vectors for Classification

Each trial is represented by the T matrix, whose dimension is $NC \times NS$, where $NC = 31$ is the number of channels (covering frontocentral to parietal areas shown in Fig. 3.3) and NS is the number of samples throughout the trial [27].

In this research segment, the entirety of the trial was taken into consideration, generating a classification sample every 100 ms. For each of these samples, the features of the classification window were computed from the preceding 1 s of data.

A sliding window with a size of WS corresponding to the time period of 1 s (preceding each time point where it is being calculated) was used to calculate the feature matrix F_i throughout the trial. The resulting F_i matrix has dimension $NC \times WS$. The feature vector $\vec{f_i}$ with NC features represents row-wise averages from the matrix F_i and is used for classification. The illustration of a feature matrix F_i for amplitude features can be seen in Fig. 5.1. Note that for spectral power features, the unit on the y axis is not μV but dB, and the number of channels is reduced (corresponding to the number of CSP features, as explained in the following subsections). Using the classification accuracy results from all subjects, the average classification accuracy is calculated for each time point i in the trial [1, 27].

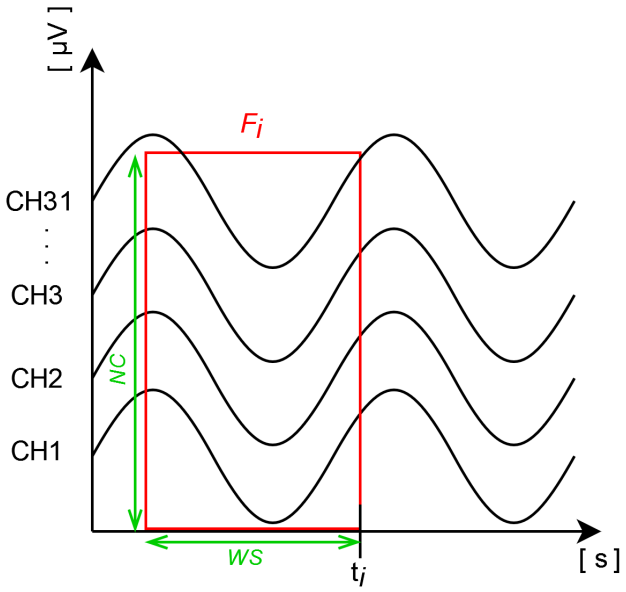


Figure 5.1: Illustration of a feature matrix F_i for amplitude features. CH1, CH2, CH3 to CH31 represent channel examples.

5.1.2. Feature Patterns for Classification

As recommended in Haufe et al. [115], activation patterns were computed for the amplitude features:

$$\mathbf{a} = \Sigma_x(\Sigma_x + \lambda\mathbf{I})^{-1}(\mu_\mathbf{1} - \mu_\mathbf{2}) \cdot Var\{\mathbf{s}\} \tag{5.3}$$

where $\mathbf{a}$ represents the activation patterns, Σ_x denotes the pooled covariance of the measurement signals $\mathbf{x}$, λ is the shrinkage parameter, $\mu_\mathbf{1}$ and $\mu_\mathbf{2}$ represent the class means, and $\mathbf{s}$ corresponds to the source estimate. The source estimate results from the backward model, i.e., the sLDA weights $\mathbf{W}$ applied to the measurements $\mathbf{x}$:

$$\mathbf{s} = \mathbf{W}^T\mathbf{x} \tag{5.4}$$

This model allows for the transformation of the input measurements into a lower-dimensional space that maximally discriminates between the classes of interest [1].

Since the spectral power features were transformed through the CSP approach, a direct depiction in topographic plots is unattainable. Therefore, rather than showcasing the classifier patterns, the projection of CSP features into the channel space has been presented. The CSP matrix $\mathbf{M}$, with each row representing a CSP filter, was computed and subsequently applied to the data $\mathbf{x}$, resulting in the acquisition of CSP features [1]:

$$\mathbf{F} = \mathbf{M}\mathbf{x} \tag{5.5}$$

Subsequently, the pseudo-inverse of the CSP model $\mathbf{M}$ was computed [106], and applied to the CSP features $\mathbf{F}$. This operation facilitated the projection of data back into the channel space:

$$\mathbf{F}_{ch} = \mathbf{M}^+\mathbf{F} \tag{5.6}$$

From there, the logarithmic power of the $\mathbf{F}_{ch}$ (relative to the baseline period) was calculated for both classes. The spectral power features' CSPs $\mathbf{P}$ were then calculated [1]:

$$\mathbf{P} = \mathbf{P}_1 - \mathbf{P}_2 \tag{5.7}$$

The spectral power feature CSPs $\mathbf{P}$ represented the differential power patterns between the classes and were used for further analysis and visualization, including the generation of topographic plots, providing insights into the spatial distribution of the discriminative power patterns across the channels and allowing for the interpretation and understanding of the classification results comparable to the above-described amplitude feature topographic plots.

To evaluate the performance of the classifications and determine their significance, a threshold of significantly better-than-chance accuracies was calculated, following the methodology described by Müller-Putz et al. [116]. This threshold served as a benchmark to assess whether the achieved accuracies were statistically significant and exceeded random chance levels. Furthermore, Wilcoxon signed-rank tests were conducted on accuracies obtained from direction classification, baseline versus MI period classification, and condition classification. This allowed for the assessment of whether there were significant differences in classification performance in this research segment [1].

5.2. Classification in Entropy Research Segment

The *Entropy research* segment, dealing with analysis of short-term entropy features, involved the classification of three distinct feature types (for both datasets): low-frequency amplitude features, entropy features with a long window (1 s), and entropy features with a short window (0.5 s) [33], as described in Section 4.1..

Each of these feature types was subjected to classification based on multiple class distributions. For the KGU dataset, the following class distributions were considered [33]:

1. Condition VtG, directions: up vs. right.
2. Condition noVtG, directions: up vs. right.
3. Condition VtG: direction right vs. baseline.
4. Condition VtG: direction up vs. baseline.
5. Condition noVtG: direction right vs. baseline.

6. Condition noVtG: direction up vs. baseline.

On the other hand, for the ULM dataset, the following class distributions were used:

1. Movements: EE vs. EF.

2. Movement EE vs. baseline.

3. Movement EF vs. baseline.

These class distributions allowed for a comprehensive analysis of the performance and discriminability of the different feature types in various experimental conditions.

For the classification of selected features in this segment of research, the sLDA classifier described in Section 5.3. was used, to remain comparable to the previous research segment.

5.2.1. Feature Vectors for Classification

Each trial is represented by the T matrix, whose dimension is $NC \times NS$, where $NC = 31$ is the number of channels (shown in Fig. 3.3) and NS is the number of samples during the MI period [27].

A sliding window with a size of WS corresponding to the time period of 0.5 s (in contrast to the previous segment of research described in Section 5.1., where the classification window was 1 s) was used to calculate the feature matrix F_i throughout the trial. A reduction in window size in this research segment was done in order to see if the results would be similar, thus reducing the potential lag during classification. The resulting F_i matrix has dimension $NC \times WS$. The feature vector $\vec{f_i}$ with NC features represents row-wise averages from the matrix F_i and is used for classification. The illustration of a feature matrix F_i for amplitude features can be seen in Fig. 5.1. Note that for entropy features, the unit on the y axis is not μV but information content (entropy). Using the classification accuracy results from all subjects, the average classification accuracy is calculated for each time point i in the trial. From this, a time point with a peak (*maximum*) of the average classification accuracy is calculated. The resulting classification accuracy for each subject is determined as the accuracy value at the said peak time point [33, 27], and the grand-average accuracy is calculated as the average of said accuracies.

5.3. Classification in Methods Comparison Research Segment

In the final segment of the research, the *Classification methods comparison research* segment, a comparative evaluation was carried out to assess the effectiveness of various classification methods, with a particular emphasis on their performance within the specific datasets investigated within the scope of this dissertation. The focus was placed on several widely used methods in MI research, including sLDA, SVM, and RF. To expand the analysis and incorporate recent advancements in deep learning, additional classifiers based on CNNs were included. Specifically, the performance of VGG-19, ResNet-101, and DenseNet-169 classifiers was assessed in the context of MI classification. Said deep learning methods were mainly built and tested for computer vision and image classification, hence, the motivation behind incorporating them in this research segment was to test if they could achieve better classification results than standard, widely used methods while interpreting EEG as an image (*samplesxchannels*). By considering a range of established and state-of-the-art methods, the research aimed to provide valuable insights into the performance and potential of different classification techniques for MI analysis for the given datasets.

5.3.1. Feature Vectors for Classification

For this part of research, only relevant MI periods were epoched for classification: from $t = 0\ s$ to $t = 2\ s$ for the KGU dataset, as shown in Fig. 3.1, and from $t = 2\ s$ to $t = 4\ s$ for the ULM dataset, as shown in Fig. 3.4 [27].

Similarly to the feature vectors described in Section 5.2., each trial is represented by the T matrix, whose dimension is $NC \times NS$, where $NC = 31$ is the number of channels (shown in Fig. 3.3) and NS is the number of samples during the MI period.

For sLDA, SVM, and RF, a sliding window with a size of WS corresponding to the time period of $0.5\ s$ was used to calculate the feature matrix F_i throughout the trial. The resulting F_i matrix has dimension $NC \times WS$. The feature vector $\vec{f_i}$ with NC features represents row-wise averages from the matrix F_i and is used for classification. The illustration of a feature matrix F_i for amplitude features can be seen in Fig. 5.1. Using the

classification accuracy results from all subjects, the average classification accuracy is calculated for each time point i in the trial. From this, a time point with a peak ($maximum$) of the average classification accuracy is calculated. The resulting classification accuracy for each subject is determined as the accuracy value at the said peak time point [27], and the grand-average accuracy is calculated as the average of said accuracies.

In the course of preliminary investigations, it was discovered that the CNN-based methods employing a window-based (WS of 0.5 s) approach yielded far more consistent accuracy within the predetermined time frame of MI movement (i.e., results were very similar at any given time point). Consequently, the decision was made to utilize the entire trial matrix T for classification with CNN models in order to reduce computational time. Due to this consistency of the CNN classifiers, no disparity was observed in the accuracy achieved between the approach using the entire trial matrix and the window-based approach [27]. In summary, the feature vector $\vec{f_i}$ with 31 components representing the mean of each EEG channel for each of the 0.5 s windows is used as input for sLDA, SVM, and RF. For CNN-based methods, the entire trial matrix T was used and interpreted as an image. It is worth noting that channels are distributed in a somatotopic manner and cannot be translated perfectly into one dimension of the feature matrix for CNN classification.

The comparison of classification methods was conducted based on the grand-average accuracy achieved by each. With respect to their implementation in the study, standard Matlab functions were employed for sLDA, SVM, and RF, whereas CNN-based methods were executed using PyTorch.

5.3.2. Shrinkage Linear Discriminant Analysis

LDA is a classification algorithm used to model the differences between groups or classes based on their input features. The main objective of LDA is to identify linear feature combinations that maximize the separation between classes. It achieves this by calculating the between-class and within-class scatter matrices, which capture the differences and similarities between classes, respectively. The goal is to maximize the between-class scatter while minimizing the within-class scatter. Once the scatter matrices are computed, LDA performs an eigenvalue decomposition to obtain the linear discriminant vectors, also

known as canonical variables or discriminant functions. These discriminant vectors define a lower-dimensional space onto which the data points are projected. In this transformed space, the classes are ideally well-separated, allowing for straightforward classification. To classify new observations using LDA, the algorithm projects the data onto the decision boundary defined by the discriminant vectors and assigns the observations to the class on the corresponding side of the boundary. This decision boundary is linear in nature, which means that LDA assumes the classes can be separated by a linear function. LDA makes certain assumptions about the input data, including the normal distribution of features within each class and the equality of covariance matrices across classes. Violations of these assumptions may affect the performance of LDA [117]. However, LDA is still widely used in various domains, such as image recognition, text classification, bioinformatics, and many other fields, due to its simplicity, interpretability, and effectiveness in many practical scenarios.

Shrinkage regularized linear discriminant analysis (sLDA) Blankertz et al. [114] is an extension of LDA designed to improve its performance in high-dimensional environments. For high-dimensional data, traditional LDA may suffer from overfitting, i.e., it may not generalize well to new data. The main idea behind shrinkage LDA is to estimate a "shrinkage" covariance matrix, which is a compromise between the sample covariance matrix and a diagonal matrix. This shrinkage matrix is computed as a linear combination of the sample covariance matrix and a diagonal matrix, with the weighting of the sample covariance matrix determined by the shrinkage parameter. By shrinking towards a diagonal matrix, the algorithm reduces the number of parameters to be estimated, which in turn reduces the risk of overfitting [27].

The classification procedure in sLDA can be summarized as follows [27]:

1. Calculate the shrinkage LDA coefficients, i.e., the weights that define the linear boundary between the different classes.

2. Calculate the class mean for each group.

3. Calculate the pooled within-class covariance matrix, which is a weighted sum of the sample covariance matrices for each group.

4. Calculate the shrinkage covariance matrix Σ_s using the equation:

$$\Sigma_s = (1 - \delta)S + \delta D, \tag{5.8}$$

where S is the sample covariance matrix, D is the diagonal matrix, and δ is the shrinkage parameter.

5. Calculate the inverse of the pooled within-class covariance matrix.

6. Calculate the discriminant value $D(x)$ for each observation using the equation:

$$D(x) = x^T \Sigma_w^{-1}(\mu_1 - \mu_2), \tag{5.9}$$

where x is the feature vector, Σ_w^{-1} is the inverse of the pooled within-class covariance matrix, and μ_1 and μ_2 are the means of the two classes.

7. Classify each observation based on the sign of the discriminant value. If $D(x) > 0$, the observation is classified as belonging to the first class. If $D(x) \leq 0$, then the observation is classified as belonging to the second class.

5.3.3. Support Vector Machine

SVM is a powerful and widely used classification algorithm that aims to find an optimal hyperplane to separate data points belonging to different classes. The key idea behind SVM is to maximize the margin, which is the distance between the hyperplane and the nearest data points of each class. In its simplest form, SVM is a binary classifier that operates in a high-dimensional space. The algorithm searches for the hyperplane that maximally separates the data points of different classes while maintaining the largest possible margin. The data points that lie closest to the hyperplane, known as support vectors, play a crucial role in determining the position and orientation of the hyperplane [118, 119]. If the data are not linearly separable, SVM uses a kernel function to map the data into a higher-dimensional space where it is more likely to be linearly separable. SVM can be used for both binary and multi-class classification tasks [27].

Kernel SVM is a variant of the SVM algorithm that addresses the challenge of nonlinearly separable data. It achieves this by mapping the input data into a higher-dimensional

feature space using a kernel function, without explicitly computing the coordinates of the transformed data points. The kernel function calculates the dot product between the feature vectors in this higher-dimensional space, allowing the SVM to operate efficiently even in high-dimensional settings. The key advantage of kernel SVM is its ability to capture complex relationships and decision boundaries in the data by implicitly transforming it into a higher-dimensional space. This is known as the "kernel trick" because it enables the SVM to effectively model nonlinear relationships without explicitly computing the transformations. By employing various kernel functions, such as linear, polynomial, radial basis function (RBF), and sigmoid, kernel SVM can adapt to different data distributions and capture diverse patterns. The linear kernel is the simplest form of kernel function and represents the original input space. It is suitable for linearly separable data. Polynomial kernels introduce polynomial terms to capture nonlinear relationships. The RBF kernel, also known as the Gaussian kernel, uses a similarity measure based on the distance between data points to model complex decision boundaries. It is effective in handling data with complex and irregular structures. The sigmoid kernel applies a sigmoid function to the dot product of the input features, allowing it to model data with a sigmoidal relationship.

During training, the SVM learns the parameters that define the hyperplane by solving an optimization problem. The algorithm tries to find the best hyperplane by minimizing the classification error and maximizing the margin. The support vectors are the data points closest to the hyperplane and play a crucial role in defining its position.

Once the SVM model is trained, it can be used to classify new, unseen data points by projecting them onto the learned hyperplane. The decision is made based on which side of the hyperplane the data point falls on.

The classification procedure in kernel SVM can be summarized as follows [27]:

1. Given a set of training data, the linear kernel SVM selects a subset of data points as support vectors, i.e., the points closest to the decision boundary in higher-dimensional space.

2. The kernel SVM then finds the hyperplane that maximizes the distance between the support vectors of each class in the higher-dimensional space.

3. To classify new observations, the kernel SVM maps them into higher-dimensional

space using the kernel function, projects them onto the hyperplane, and assigns them to the class on the corresponding side. The sign of the projection determines the class of the observation.

5.3.4. Random Forest

The RF method uses a group of predictors (e.g., decision trees) in order to aggregate the prediction votes and then predict the outcome that gets the most votes. The group of predictors is called an ensemble, thus, methods such as RF are called ensemble methods. In the present research, RF with 20 decision trees (DTs) is utilized [120, 119]. In this method, multiple decision trees are built using a random subset of the training data and a random subset of the input features. The final prediction is made by averaging the predictions of all the trees in the forest. RF has several advantages that make it well-suited for handling high-dimensional data and noisy datasets, such as EEG. Firstly, it reduces the risk of overfitting by building multiple trees with different subsets of data and features, which helps capture different aspects of the data's variability. Secondly, RF can handle a large number of input features without feature selection or dimensionality reduction techniques, making it suitable for large datasets. Additionally, RF has built-in mechanisms to assess the importance of features, providing insights into their relevance for the classification task. The method works as follows [27]:

1. Randomly sample the training data with replacement (bootstrap) to create multiple datasets (or decision trees) of the same size as the original dataset.

2. For each dataset, randomly select a subset of the input features to use for building the DT.

3. Build a DT for each dataset using the selected features and a splitting criterion.

4. Repeat steps 1-3 to create a forest of DT.

5. To make a prediction for a new sample, pass it through all the DTs in the forest and average their predictions (for regression tasks) or take the majority vote (for classification tasks).

One way to obtain a diverse set of classifiers in ensemble methods (e.g., RF) is to use very different training methods. Another approach is to use the same training method for each predictor but train them on different random subsets of the training set. When sampling is performed with replacement, this method is called bagging. Bagging allows training instances to be sampled several times across multiple predictors. Bagging is utilized in RF implementation in the present research.

5.3.5. VGG-19

The VGG-19 architecture [121] is a deep CNN known for its simplicity, depth, and strong performance in various computer vision tasks such as image classification and object detection.

VGG-19 consists of 19 layers: 16 convolutional layers and 3 fully connected layers. These layers are organized into 5 convolutional blocks and a fully connected block. Each convolutional block contains a varying number of convolutional layers with small 3×3 filters, followed by a max-pooling layer. The fully connected block consists of 3 fully connected layers, followed by a softmax activation function [122] to output class probabilities.

Convolution allows the model to learn local patterns and spatial features in the data by sliding a filter K over the input data, computing an element-wise multiplication between the filter and the input, and then summing the results. The convolution operation can be expressed as follows [27]:

$$F(m,n) = \sum_i \sum_j I(m-i, n-j) \cdot K(i,j), \tag{5.10}$$

where F is the feature map output, I is the input, K is the filter/kernel, and (m,n) and (i,j) are indices for the spatial dimensions of the output feature map and the kernel, respectively. The convolution operation is applied with a stride of 1 and a padding of 1 to preserve the spatial dimensions of the input.

To reduce the spatial dimensions of the feature maps while preserving the most important information, the max-pooling operation [123] is performed on the output features.

The fully connected layers in VGG-19 are used to combine the features learned from the previous layers by performing matrix multiplication between the input data and the weight matrix, followed by the addition of a bias term:

$$Y = W \cdot X + b, \tag{5.11}$$

where Y is the output, W is the weights matrix, X is the input, and b is the bias term.

Finally, the output of the last fully connected layer is passed through a softmax activation function to produce class probabilities.

The VGG-19 architecture is trained using a backpropagation algorithm to minimize a cross-entropy loss function:

$$L(y, \hat{y}) = -\sum_{i=1}^{N} y_i \log(\hat{y}_i), \tag{5.12}$$

where $L(y, \hat{y})$ represents the cross-entropy loss between the true and the predicted probability distribution, and N is the total number of classes.

The weights and biases of the network are updated by the gradient descent optimization algorithm Adam [124], with $lr = 0.0001$. By employing small filter sizes, deep architecture, and consistent design principles, VGG-19 has demonstrated exceptional performance in large-scale image recognition tasks.

5.3.6. ResNet-101

ResNet-101 is a specific instantiation of Residual Networks (ResNets) [125], a family of deep CNNs. ResNet-101 consists of 101 layers, including convolutional layers, batch normalization layers, activation layers, and pooling layers, as well as residual connections.

A residual block presents the basis of ResNet architecture and can be represented as [27]:

$$y = F(x, W) + x, \tag{5.13}$$

where y is the output of the residual block, x is the input to the block, $F(x, W)$ represents the residual operations applied, and W denotes the set of learnable weights associated with the block. The addition operation between $F(x, W)$ and x forms a shortcut connection, allowing the gradients to flow more effectively during backpropagation, thus improving the optimization process [27].

ResNet-101 is structured as a series of stacked residual blocks, with each block contain-

ing multiple convolutional layers. The first few layers perform initial feature extraction, while subsequent layers capture increasingly complex patterns. ResNet-101 consists of four groups of residual blocks, with each stage featuring a different number of blocks and output feature map dimensions. Down-sampling occurs between stages through pooling operations [27].

The model concludes with a global average pooling layer (GAP). For an input feature map F with dimensions $H \times W$, the GAP operation can be defined as:

$$GAP(F) = \frac{1}{H \times W} \sum_{i}^{H} \sum_{j}^{W} F(i, j), \tag{5.14}$$

where $F(i, j)$ is the value at position (i, j) of the input feature map.

The GAP layer is followed by a fully connected layer that outputs class probabilities via a softmax activation function.

The goal of training ResNet-101 is to minimize a loss function, in this case a cross-entropy loss, using the Adam gradient descent optimizer with $lr = 0.0001$. By employing residual learning and deep architecture, ResNet-101 has demonstrated exceptional performance on various computer vision tasks, including image classification and object detection, outperforming shallower CNN architectures and other classification methods not based on deep learning [27].

5.3.7. DenseNet-169

DenseNet-169 is a deep CNN architecture that belongs to the dense convolutional networks (DenseNets) family [126]. DenseNet-169 is particularly known for its efficient use of network parameters and its ability to alleviate the vanishing gradient problem, which enables the training of deep architectures while maintaining strong performance in various computer vision tasks such as image classification and object detection.

The defining characteristic of DenseNets is the dense connectivity pattern, where each layer receives the feature maps from all preceding layers as input. This is in contrast to conventional CNNs, where each layer receives the inputs of its immediate predecessors. Mathematically, the output of the l-th layer in a DenseNet can be represented as follows [27]:

$$H_l = f_l([H_1, H_2, \ldots, H_{l-1}]), \tag{5.15}$$

where H_l is the output feature map of the l-th layer, f_l is the layer function (e.g., convolution, activation function), and $[H_1, H_2, \ldots, H_{l-1}]$ denotes the concatenation of feature maps from layers 1 to $(l-1)$.

DenseNet-169 consists of 169 layers and four dense blocks with varying numbers of densely connected layers and three transition layers interspersed between the dense blocks. The network starts with an initial convolutional layer, followed by dense blocks and transition layers in alternating order. The last dense block is followed by the GAP layer, which aggregates the feature maps into a compact representation, and a fully connected layer with a softmax activation function to generate class probabilities as the final output [27].

Like VGG-19 and ResNet-101, the DenseNet-169 architecture is also trained with the backpropagation algorithm to minimize the cross-entropy loss function, updating the weights and biases of the network by the gradient descent optimization algorithm Adam, with $lr = 0.0001$. The dense connectivity pattern in DenseNet-169 facilitates efficient gradient flow and feature reuse, resulting in improved performance with fewer parameters compared to conventional deep CNN architectures [27].

Chapter 6

RESULTS AND DISCUSSION

This chapter presents, interprets, and discusses the results of the research conducted as part of this dissertation. The chapter is divided into sections that correspond to the three segments of the study and contextualize the contributions of the research, as outlined in Section 1.5..

Section 6.1. outlines the results obtained during the first segment of the study, the *KGU research*, focusing on the neurophysiology of the newly curated data, initial MI classification results, and behavioral outcomes.

Section 6.2. addresses the outcomes of the second segment of the research, the *Entropy research*, which focuses on the short-term information entropy based on TFRs and their effectiveness in successfully detecting MI detection.

Finally, Section 6.3. presents and interprets the results of the third segment of the research, the *Classification methods comparison research*. In particular, this involves a statistical investigation of the accuracy comparison of the classifiers described in Section 5.3., followed by an assessment of these classifiers.

6.1. KGU Research Segment Results

In this section, the comprehensive results obtained as part of the *KGU research* segment are presented, analyzed, and discussed. The important contributions to the understanding of the properties of MI in the new context of vibrotactile-guided paradigms are brought forth by the comprehensive insights generated in this segment. The dynamics

of the human brain in such scenarios are demonstrated through insightful observations regarding the nuances of neurophysiological patterns that unfold during MI tasks with different guidances. The sLDA classification of MI directions is explored, and the discriminability of MI versus baseline states is compared, shedding light on the intricate neural patterns that render these tasks distinguishable. Furthermore, the impact of vibrotactile guidance on MI task performance is investigated, offering novel perspectives on the utility of such tactile stimulation. Lastly, the behavioral outcomes from the questionnaire analysis are presented and discussed to provide critical insight into the real-world implications and potential applications of the results presented.

6.1.1. Implications for Neurophysiology

In the following subsections, the dynamics of the human brain in a paradigm with visual or a combination of visual and vibrotactile guidance are comprehensively examined, and the implications of these dynamics for neurophysiology are analyzed. Amplitude potentials are examined first, followed by a time-frequency analysis.

6.1.1.1. Amplitude Potentials

Fig. 6.1 illustrates the amplitude potentials between 0.2 and 5 Hz during an average trial. The potentials are shown separately for each condition and each direction. Potentials are presented as 95% confidence intervals of the amplitudes at Pz, CPz, and Cz, supplemented by topographic diagrams at selected time points. The left panel shows the potentials for the condition VtG, while the dark green trace corresponds to the direction *right* and the light green trace to the direction *up*. Similarly, for the condition noVtG in the left panel, the potentials for the direction *right* are shown in dark blue and the potentials for the direction *up* in light blue. For the time points before the start of the cue motion, one set of topographic diagrams is presented per condition (left and right panels), while for the time points after the start of the cue motion, two sets are presented in each panel, one for each direction. These diagrams are framed with a color-coded border corresponding to the color of the amplitude traces of the corresponding direction (dark shade green/blue is direction *right*; light shade green/blue is direction *up*). The sketch in the center below highlights the electrode positions whose amplitude traces are shown [1].

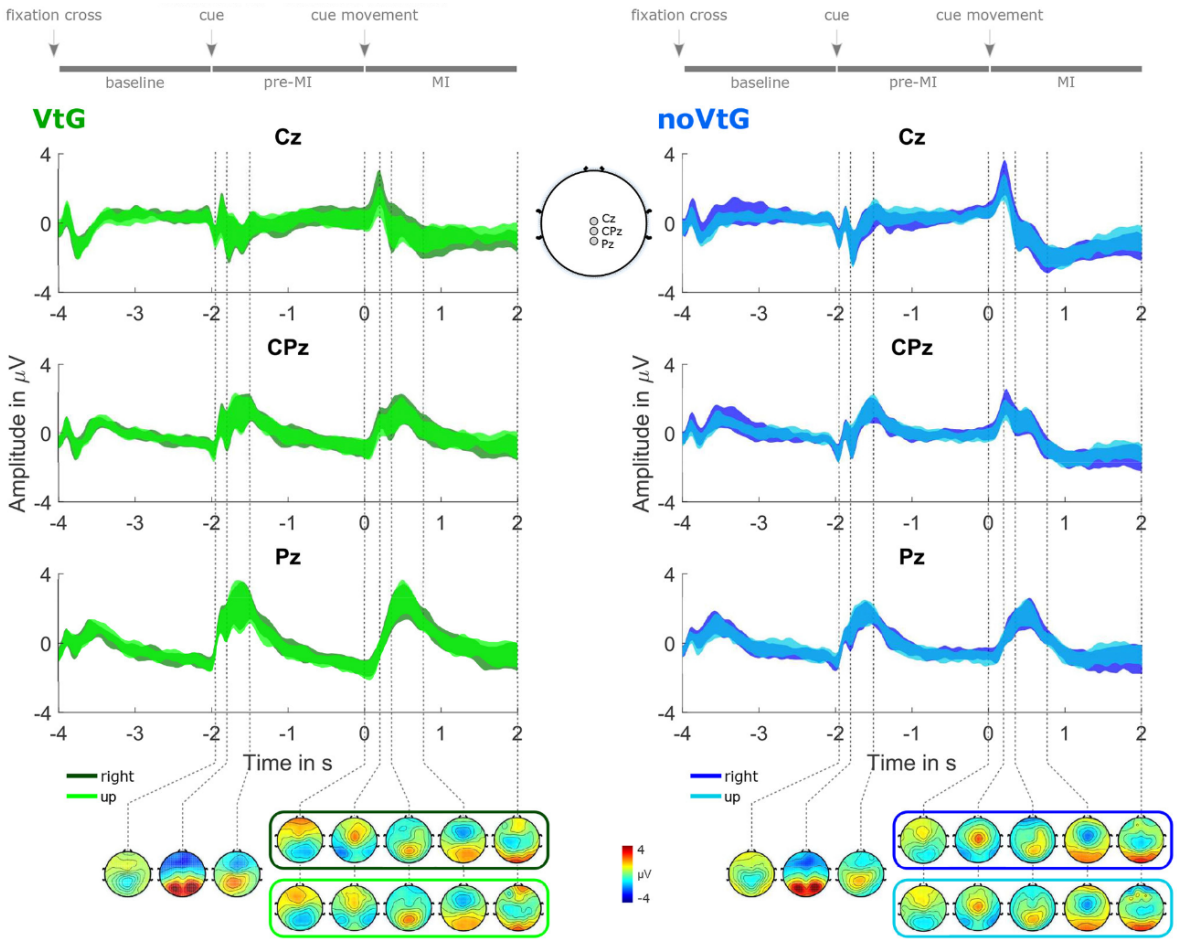


Figure 6.1: Grand-average potentials ($0.2 - 5\ Hz$), for each condition, and direction [1].

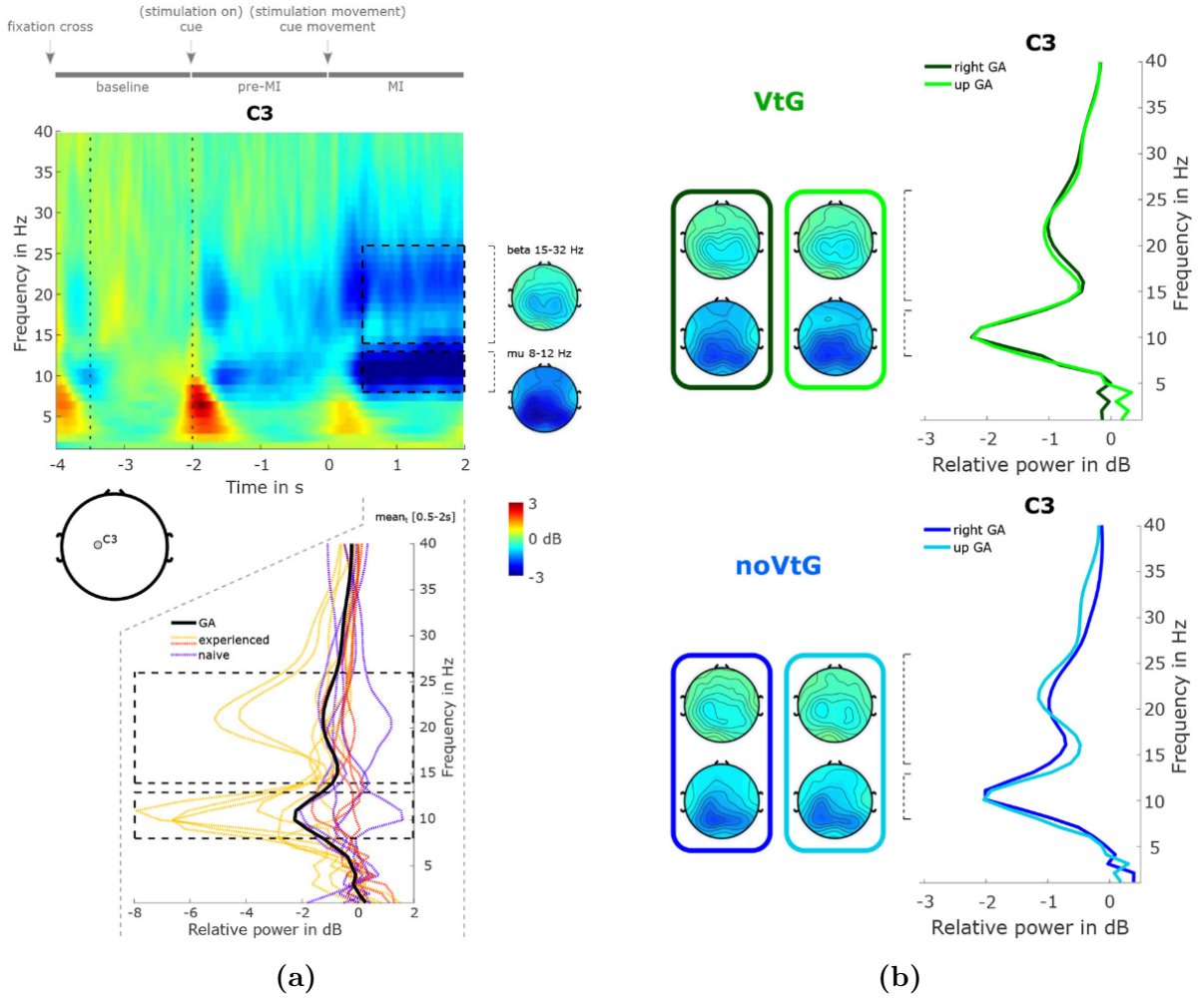
There are prominent evoked responses to the stimuli provided by the paradigm, as shown in Fig. 6.1. Visual evoked potentials (VEPs) are present in both conditions, following the appearance of the fixation cross, the appearance of the visual cue, and the onset of the cue motion, with the VEP evoked by the fixation cross being smaller than the other two. In the VtG condition, the second VEP overlaps with a somatosensory evoked potential (SEP) evoked by the onset of the vibrotactile stimulation. During the period of MI, an MRCP can be observed that presents as a central negativity that peaks within 1 s of cue motion onset and partially overlaps with the VEP. The peak amplitudes, slopes, and spatial profiles of the VEP and the MRCP vary slightly with the movement direction and the condition. The MRCP negativity is stronger and less variable in the noVtG condition (peak mean $\pm$ std at $t = 0.77$ s: $-0.60 \pm 1.71\mu$ V in VtG, $-1.83 \pm 1.79\mu$ V in noVtG), but spatially broader in the VtG condition. In both conditions, it is initially precentral and later more centrally localized, with the later component showing contralateralization in the VtG condition but not in the noVtG condition [1].

The grand-average potentials shown in Fig. 6.1 indicate slight differences in the positive peak amplitude of the evoked potentials, caused by the onset of cue motion between directions, with the positive peak being stronger for trials in the *right* direction in both conditions. It is possible that the inherent internal mapping between the cue moving upward and the imagined movement forward/to the front may influence the intensity of the potential. The positive peak is followed by a combination of two negative peaks, the first of which is attributed to the late negative component of the evoked potential, while the second vaguely exhibits the characteristics of an MRCP as the central negativity expected when the imagined movement is initiated. Because the onset of the MRCP depends on exactly when the imagining is initiated and thus varies between individual subjects and individual trials, it is not surprising that there is considerable variance in the grand-average potential. The spatial profile of the MRCP appears to be influenced by the vibrotactile guidance, being more contralateral in the VtG [1].

6.1.1.2. Time-frequency Analysis

Fig. 6.2a shows a time-frequency map of the grand-average trial (both conditions), along with topographic diagrams of the mu and beta bands and the relative power spectrum for the MI period. The power spectrum diagram includes the grand-average and

the individual power spectra [1].



(a) **(b)**

Figure 6.2: Overview of spectral power features. 6.2a: Time-frequency decomposition of the grand-average trial, topographic diagrams, time-frequency map, and power spectrum. 6.2b: Power spectra at C3 and topographic diagrams for conditions VtG and noVtG and directions *right* and *up* [1].

Here, the time-frequency decomposition and spectrum are calculated with respect to the period marked by dashed vertical lines in the time-frequency map (-3.5 to -2). The time-frequency map and spectrum are shown at position C3. The topographic diagrams were calculated from the regions marked by the black dashed rectangles in the time-frequency map, i.e., $t = 0.5\ s$ to $2\ s$ in the time dimension and $8-12\ Hz$ for the mu band and $15-32\ Hz$ for the beta band in the frequency dimension. The bottom panel shows the grand-average (black solid line) or the spectra of the individual subjects (colored dashed lines), respectively, during the MI period ($t = 0.5\ s$ to $2\ s$). The individual spectra are divided into three subgroups, namely MI-experienced with above-average spectral

peaks (yellow), MI-experienced with average or weaker mu peaks (red), and MI-naïve (purple). In Fig. 6.2b, the colors green and blue denote the conditions. The spectra and topographic diagrams were calculated from the same areas shown in Fig. 6.2a [1]. Here, a decrease in power in the mu and beta frequency ranges during the MI period over the centro-parietal areas can be seen, and a weaker decrease during the period before MI. As can be seen from the individual relative power spectra in Fig. 6.2a (bottom panel), there is considerable variance between subjects. In the group of participants with prior MI experience, six subjects showed above-average desynchronization (yellow, mu peaks mean $\pm$ std: $-5.52 \pm 2.18\ dB$) and four showed weaker-than-average desynchronization (red, mu peaks mean $\pm$ std: $-1.26 \pm 1.00\ dB$). In the group of participants with no prior MI experience, all five individuals showed average or weaker-than-average desynchronization (purple, mu peaks mean $\pm$ std: $-1.24 \pm 0.92\ dB$). In Fig. 6.2b, the grand-average spectra are shown separately for each condition and each direction. The spectral profiles are very similar between the conditions and the directions, with the mu peak being slightly stronger in the VtG condition ($-2.25\ dB$ vs. $-2.04\ dB$). Topographically, the strongest decrease is located over centro-parietal areas in both conditions, while it is more lateralized to the contralateral hemisphere in the noVtG condition [1].

As for the decrease in mu and beta power during the MI period shown in Fig. 6.2a, there is considerable variability between subjects, with three subjects showing particularly strong power patterns and five subjects showing very weak or uncharacteristic power patterns. Some of this variability may be explained by the degree of experience with MI tasks. Indeed, all six participants with above-average mu and beta peaks had previous MI experience, and the five MI naïve participants are among the eight with weakest mu and beta peaks. However, it is well documented that ERD/ERS profiles in general can vary considerably across individuals [127, 128, 129, 130], and while MI is a skill that can be trained and refined, naïve individuals may exhibit weak or strong power patterns based on their neuroanatomy. The separation in the strength of the power profiles between the three groups shown in Fig.6.2a is to some extent also visible in the classification results of MI versus baseline based on spectral power features in Fig. 6.5, especially in the VtG condition. Between conditions and directions, the grand-average spectral profiles show small variations. The mu peaks are fairly consistent, especially for experienced subjects, and the beta peaks vary slightly in strength, with the peaks for the *up* direction being

stronger. Spatially, the pattern under the VtG condition tends to be bilateral compared to the noVtG condition, where it is mostly contralateral. In addition, the pattern under the VtG condition is somewhat broader [1].

It is worth noting that the condition VtG contained an additional task, namely to answer after each trial whether the two guidance modalities were congruent, as mentioned in Section 3.1.2.. However, this task was performed outside the main trial period and with sufficient temporal spacing to avoid contamination of the signals of interest by the additional motor activity of the key press motion. On average, the key press occurred 2.4 s after the end of the trial, and 5 to 6 s passed thereafter until the start of the baseline of the next trial. To the best of the author's knowledge, these time intervals are sufficient to avoid overlap of the signals of interest with ERD [20] or an MRCP [23] triggered by the key press motion [1].

6.1.2. Directions Classification

Accuracies for sLDA classification between directions based on amplitude features are shown in Fig. 6.3. The grand-average accuracies are shown as thick solid lines, in green for the VtG condition and in blue for the noVtG condition. Individual subject accuracies are represented by dashed dark gray lines, and individual peak accuracies are marked with squares for MI experienced participants and diamonds for MI naïve participants, with the same color coding as in Fig. 6.2. Below the accuracy diagrams, the activation patterns for selected time intervals during the MI period are shown [1].

The average peak accuracy, which is the mean of the subject-specific peak accuracies, is 69.67% in the VtG condition and 67.01% in the noVtG condition. Subject-specific peak accuracies are not significantly different between the two conditions when using a Wilcoxon signed-rank test ($p = 0.1354$). In both conditions, activation patterns are strongest between 0.6 and 1.2 s after the onset of cue motion (positive), where most individuals' accuracies peak, and at the end of the trial (negative). Spatially, activation in the VtG condition is strongest in central and frontocentral areas, whereas in the noVtG condition it is concentrated in central and parietal areas [1].

The sLDA classification between directions based on spectral power features can be seen in Fig. 6.4. For both conditions, grand-average accuracy is within the chance level

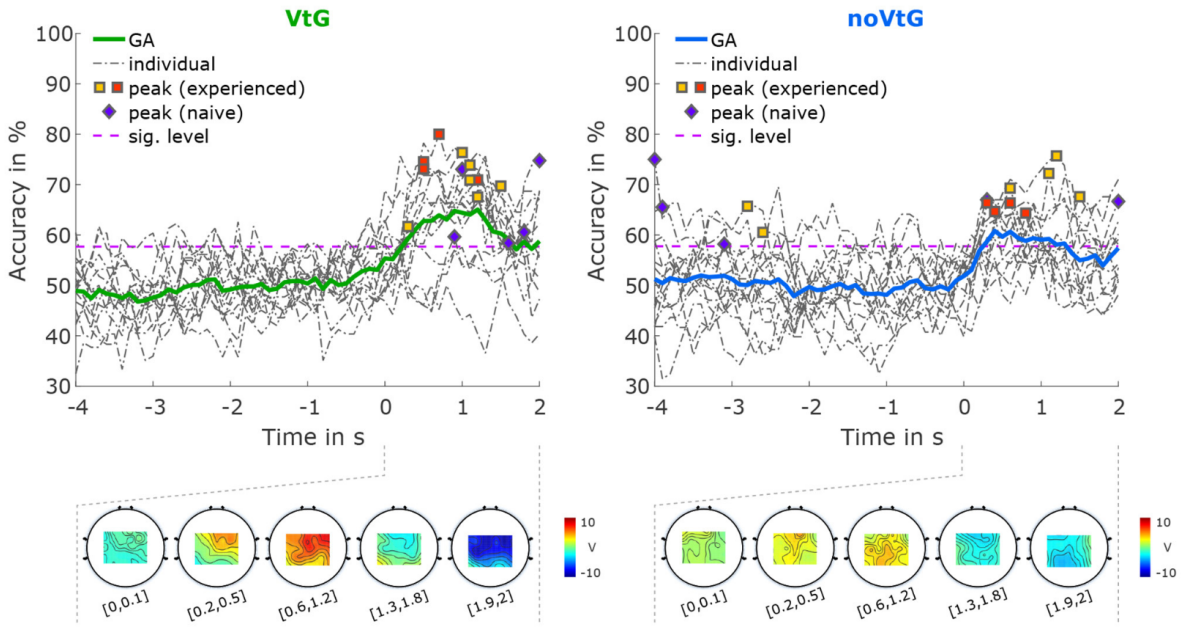


Figure 6.3: sLDA Classification results diagrams (for each condition separately) for direction *right* vs. direction *up*, based on amplitude features [1].

and does not exceed the significance level [1].

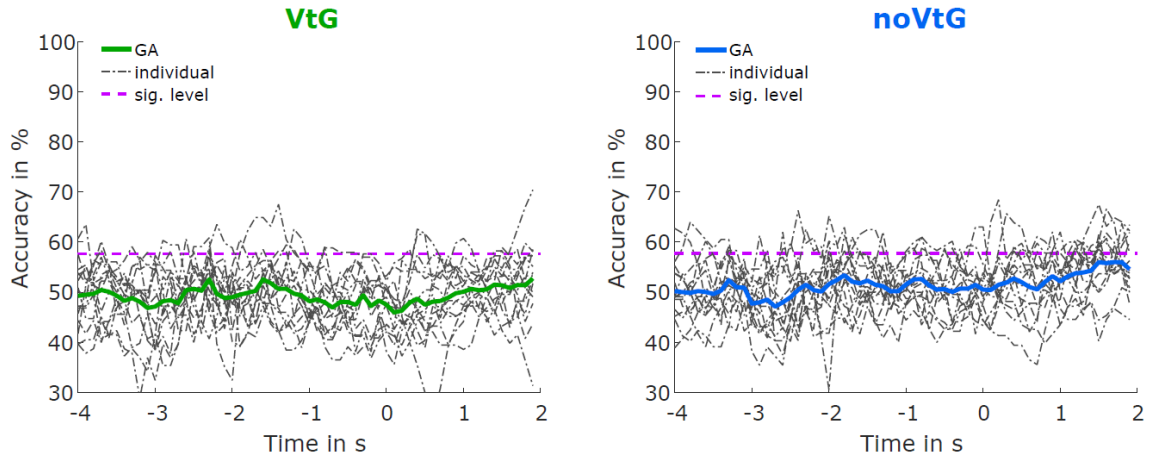


Figure 6.4: sLDA Classification results diagrams (for each condition separately) for direction *right* vs. direction *up*, based on spectral power features [1].

For classification between directions based on amplitude features (Fig. 6.3), accuracies were above the significance threshold during the MI period, with peak accuracies of $58-80\%$ (grand-average 65%) in the VtG condition and $58-76\%$ (grand-average 61%) in the noVtG condition. While the difference in peak accuracies shown here is not significant

according to the Wilcoxon signed-rank test, the accuracies in the VtG condition are higher in absolute numbers and show less variability between subjects. In addition, condition VtG exhibits stronger activation patterns that appear to be a mixture of central and parietal activations. In contrast, the activation patterns in condition noVtG are more centered in the postcentral and parietal areas, but are strongest parietally. Kobler et al. [22] found that directional information is encoded in parietal areas as well as, to a lesser extent, in central motor areas. Apparently, the classification results presented here are the result of a combination of both networks, although motor areas seem to be more represented in the VtG condition. It is worth noting that Kobler et al. [22] concluded that decoding performance is not driven by differences in evoked potentials [1].

Although the between-direction classification based on spectral power features (Fig. 6.4) did not yield accuracies that exceeded the significance level for either condition in this segment of research, this is to be expected because the spectral power features are broadly distributed across the sensorimotor cortex. Therefore, they may not be sensitive enough to the slight spatial variations in brain activity that could correspond to different imagined directions of movement of the same limb. Overall, the inability to exceed significance in the classification accuracy could suggest that more spatially precise or temporally sensitive features obtained through different spatial reduction methods such as Filter Bank Common Spatial Pattern combined with Stationary Subspace Analysis might be necessary to effectively distinguish the direction of MI from such features [131].

6.1.3. Motor Imagery vs. Baseline Classification

Fig. 6.5 shows the distribution of accuracies for the sLDA classification of MI versus baseline, separated by condition. The distributions of accuracies are presented as box plots, with individual accuracies labeled with the same color coding as in Figures 6.2 and 6.3. Statistical differences are marked with an asterisk. The activation patterns of the amplitude features and the spectral power features (CSPs) are shown below the box plots.

Peak accuracies of over 90% were observed for both amplitude and spectral power features under both conditions. The minimal accuracy varies from 58.8% (noVtG condition, spectral power features) to 75.7% (VtG condition, amplitude features). Average accuracies were found to be higher under the VtG condition (amplitude 83.2%, spectral 82.6%)

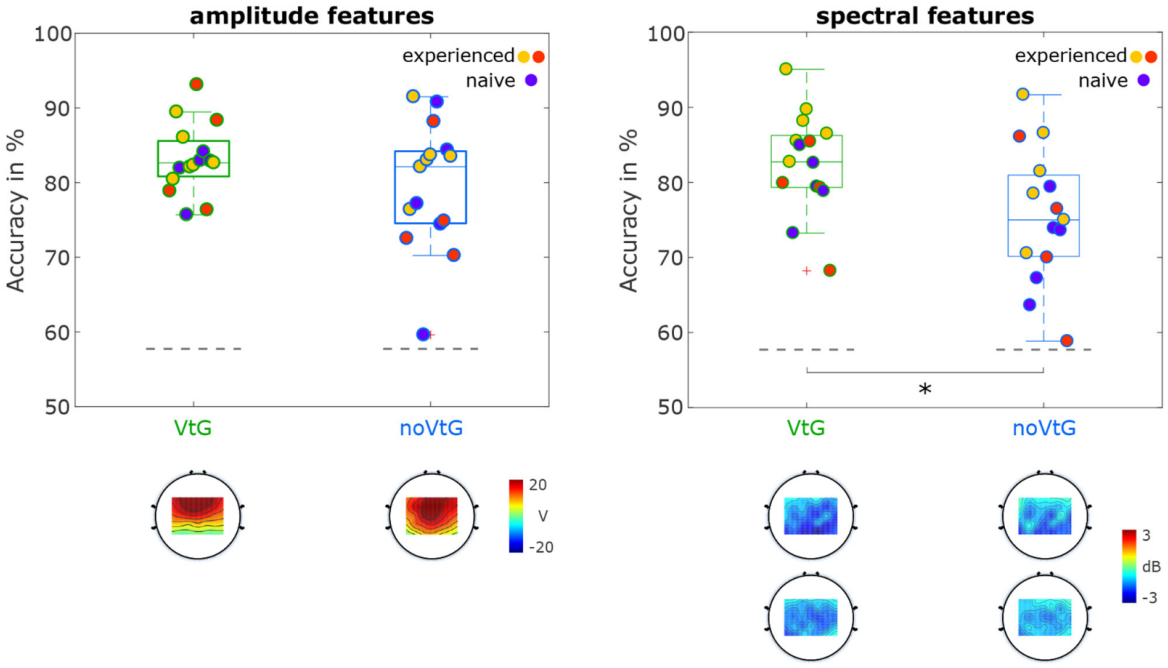


Figure 6.5: sLDA Classification results diagrams for MI vs. *base*, for each type of feature and each condition separately [1].

than the noVtG condition (amplitude 79.5%, spectral 75.5%) for both feature types, and less variance in individual accuracies was observed under the VtG condition. A significant discrepancy in individual accuracies ($p = 0.0012$) was revealed by the Wilcoxon signed-rank test for spectral power features, whereas no significant discrepancy was observed for amplitude features ($p = 0.0730$). The outcome of the significance tests is mirrored in the disparity in the medians, as highlighted in the box plots. For spectral power features, the median under the VtG condition (82.7%) exceeded that of the noVtG condition (75.0%), while the medians for amplitude features were almost the same (VtG 82.6%, noVtG 82.1%).

In the noVtG condition, the activation patterns for the amplitude features were primarily located in central channels, and they were found to be slightly more frontal under the VtG condition, with similar intensities in both. The spectral power features' CSPs were mainly found in the parietal and central regions and showed stronger intensity in the mu frequency band than in the beta frequency band. Spatially, similar patterns were observed under both conditions, though a slight enhancement was detected in the VtG condition [1].

In the classification of the MI periods against the baselines, as illustrated in Fig. 6.5, the individual exhibiting the highest performance under each condition and for each feature type achieved accuracies that surpassed 90%. For both feature types, improved results, specifically higher mean accuracies and lower accuracy variance, were observed under the VtG condition. However, according to the Wilcoxon signed-rank test, the medians exhibit significant differences only in the case of the spectral power features. It is demonstrated by the spectral power features' CSPs that there is a power reduction during the MI period relative to the baseline, in both mu and beta frequency bands, across parietal and central regions under both conditions. The reduction was found to be more pronounced in the mu frequency band under both conditions. The power reduction was more significant (across both frequency bands) under the VtG condition compared to the noVtG condition, suggesting that the vibrotactile guidance exerts a greater influence on the power difference between the MI period and the baseline, with the most significant impact occurring in the mu frequency band. Conversely, the activation patterns of the amplitude features exhibit a slight variance in their spatial distribution, where the pattern under the VtG condition is more prominently centered in the frontal regions. These patterns mirror the spatial profiles of the potentials within the time window employed for these classifications [1].

6.1.4. Conditions Classification

Fig. 6.6 presents the sLDA classification accuracies for comparisons between the two conditions (VtG and noVtG). Grand-average accuracies are represented by thick black lines, while single-subject accuracies are depicted by dashed dark gray lines. Squares denote the individual peak accuracies for participants experienced in MI, and diamonds indicate the individual peak accuracies for MI naïve participants, with the same color scheme used in Figures 6.2, 6.3, and 6.5. The activation patterns of the amplitude features (bottom left) and the spectral power features' CSPs (bottom right) are illustrated for selected time intervals within the pre-MI and MI periods [1].

With the utilization of amplitude features, the average peak accuracy reached was 80.39% (std 5.41%). Accuracy levels surpassing 70% were attained during both the MI period and the pre-MI period, with a maximum grand-average accuracy of 75.00%. Re-

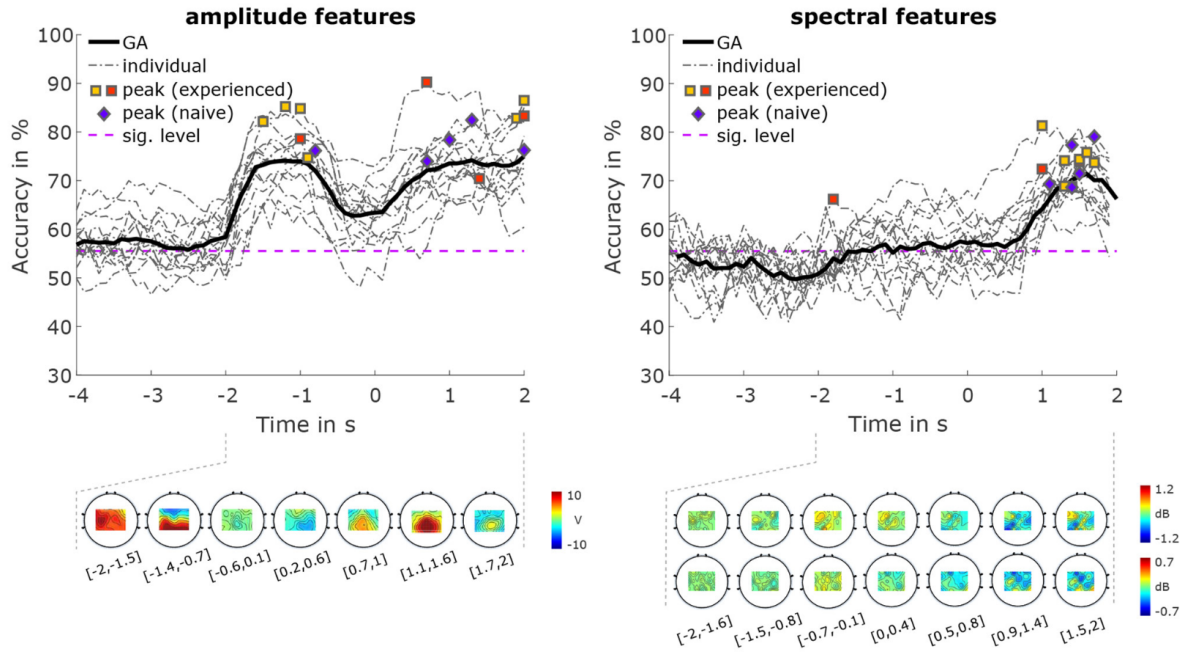


Figure 6.6: sLDA Classification results diagrams (for each type of feature separately) for condition VtG vs. condition noVtG [1].

garding spectral power features, an average peak accuracy of 73.04% (std 4.25%), and a maximum grand-average accuracy of 71.4% were achieved. The amplitude features' activation pattern is most distinct between $t = -1.4\ s$ and $-0.7\ s$ during the pre-MI period. This pattern shows a strong positive inclination in parietal regions and a negative tendency in frontal areas. During the MI period, the patterns exhibit robust positivity in the central regions, peaking between $t = 1.1\ s$ and $1.6\ s$.

The spectral power features' CSPs for the mu frequency band (Fig. 6.6) reveal that there is a more substantial negativity (indicative of a greater power reduction under the VtG condition than the noVtG condition) in parietal regions during the MI period (specifically, from $t = 0.5\ s$ to $2\ s$). Concurrently, a more pronounced positivity (implying a greater power reduction under the noVtG condition than the VtG condition) is visible in the frontal areas during the same time period. During the pre-MI period (specifically, from $t = -0.7\ s$ to $0.4\ s$), there is prominent positivity (suggesting a more substantial power reduction under the VtG condition than the noVtG condition) in frontal regions.

For the beta frequency band during the MI period (particularly from $t = 0.5\ s$ to $2\ s$), spectral power features' CSPs demonstrate greater negativity (implying a more significant power decrease under the VtG condition than the noVtG condition) in central motor

regions, and more pronounced positivity in posterior parietal regions. During the pre-MI period (specifically, from $t = -0.7$ s to 0.1 s), a more noticeable positivity (indicative of a greater power decrease under the noVtG condition than the VtG condition) is observable in central and parietal regions [1].

The classification between conditions, as displayed in Fig. 6.6, indicates that signals can be adequately distinguished during the MI period, with grand-average accuracies exceeding 70% being achieved by both feature types. For amplitude features, a similar level of performance is observed during the pre-MI period, with some participants even recording higher peak accuracy during the pre-MI period compared to the MI period. However, considering the activation patterns, the ability to discriminate between the two intervals is attributed to distinct underlying differences. During the MI period, the pattern focuses on central and postcentral areas, while the pre-MI period pattern comprises a positive parietal activation and a negative frontal activation. In contrast, for spectral power features, accuracy during the pre-MI period only marginally surpasses the baseline and barely exceeds the significance threshold (peak grand-average 57%). This insinuates that vibrotactile stimulation does not necessarily induce ERD in the absence of a motor task, thereby bolstering a previous observation from Hehenberger et al. [35], where no ERD was identified in a non-movement condition featuring vibrotactile stimulation. Spectral power features' CSPs demonstrate that a power decrease during the MI period occurs in both the mu and beta frequency bands. In the mu band, it is more parietal, while in the beta band, it is more central. Given that power changes occur simultaneously (during MI) across both frequency bands but are observed in different topological regions, this could provide a rationale for why the most successful classification results were obtained when both frequency bands were combined. It also suggests that both the mu and beta bands carry significant information about power changes for the proposed feature classification [1].

6.1.5. Questionnaire Analysis

Table 6.1 presents a summary of the participants' responses to chosen questions from the questionnaire, and it is supplemented by Fig. 6.7. Two participants didn't complete the questionnaire properly (one entirely, one partially), and their missing responses are

represented by question marks in both Table 6.1 and Fig. 6.7.

Each question could be responded to by assigning a rating on a Likert scale from 1 to 5. The top rows in Table 6.1 demonstrate the ratings concerning the perceived mental and physical tiredness of the task, where a lower rating implies it was not exhausting, while a higher rating suggests it was very tiring. The bottom rows exhibit the ratings for how easy the participants found it to maintain focus throughout the experiment, to concentrate on the MI task, and for condition VtG, how easy they found it to concentrate on the vibrotactile guidance. In summary, a high rating for the top two questions signifies high effort (very tiring), while a high rating for the bottom three questions represents low effort (easy to remain focused/concentrate). The full questionnaire can be found in Appendix A..

The columns labeled 1-5 (for each condition) comprise the number of participants who gave the respective rating on the questionnaire. The columns marked with a question mark denote the number of participants who failed to correctly complete the questionnaire, and the far-right column for each condition displays the average rating for each question. The two far-right columns of the table indicate the count of participants who provided a higher rating to the respective questions in either condition [1].

In Fig. 6.7, darker hues correlate to a lower rated effort, while lighter tones denote a higher rated effort (for instance, "not at all tiring" or "very easy to concentrate"). The top row (green colors) pertains to condition VtG, while the bottom row (blue colors) refers to condition noVtG.

Table 6.1: Overview of subjective ratings on the questionnaire [1].

	VtG							noVtG							more in VtG	more in noVtG
	1	2	3	4	5	?	avg	1	2	3	4	5	?	avg		
mentally tiring	2	2	2	5	2	2	3.2	2	4	1	6	1	1	3.0	3	1
physically tiring	8	2	2	1	0	2	1.7	7	3	1	3	0	1	2.0	0	3
able to remain focused	0	3	1	6	3	2	3.7	0	4	3	4	3	1	3.4	2	1
easy to concentrate on MI	0	2	4	3	4	2	3.7	0	1	2	5	6	1	4.1	3	5
easy to concentrate on VG	0	1	2	2	8	2	4.3	-	-	-	-	-	-	-	-	-

Seven participants rated each condition as mentally exhausting (rating 4 or 5). For condition VtG, three participants found it more tiring than condition noVtG, while one participant found the reverse to be true. In terms of physical exertion, each condition was rated as not tiring (rating 1 or 2) by ten participants, with one person finding condition

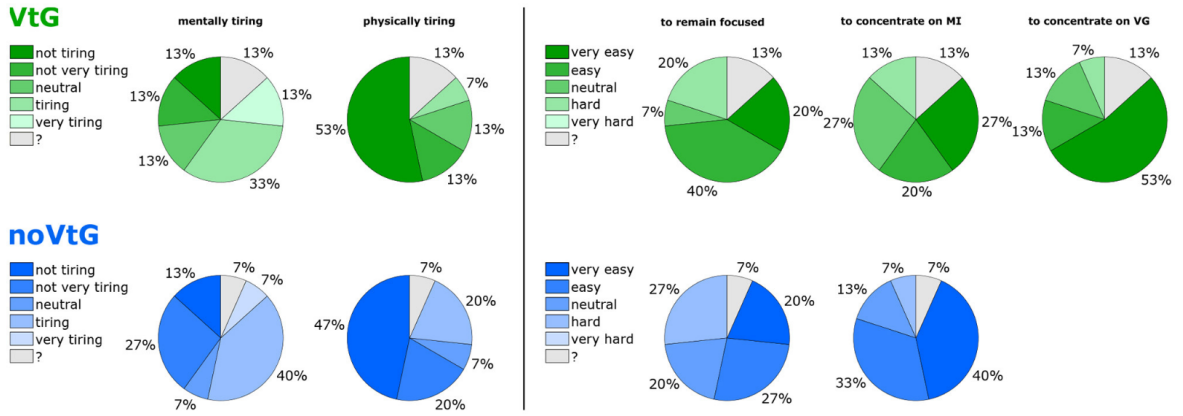


Figure 6.7: Overview of subjective participants' ratings on the questionnaire [1].

noVtG more exhausting, and none finding VtG more so. Notably, no participant assigned a rating of 5 in this category for either condition [1].

Regarding focus maintenance throughout the experiment, nine participants expressed their ease (rating 4 or 5) for condition VtG, while seven did so for condition noVtG. Two participants assigned a higher rating for condition VtG and one participant for condition noVtG. In terms of ease in focusing on the MI task (rating 4 or 5), seven participants found it easy in condition VtG and eleven in condition noVtG. Out of these, three found it simpler in condition VtG and five found it simpler in condition noVtG. It was noted that ten participants found it straightforward to focus on the vibrotactile guidance (rating 4 or 5) [1].

Additionally, all participants demonstrated high or very high confidence that they identified the majority of the incongruent trials. On average, participants accurately detected 10.07 (std 1.94) out of 12 incongruent trials and incorrectly marked 0.63 (std 1.25) out of 120 regular (congruent) trials as incongruent. This considerable agreement in detection verifies that participants were indeed successful in remaining focused during the task. Table 6.2 presents the count of trials mistakenly identified as incongruent (false positives) and the count of trials accurately identified as incongruent (true positives) by the participants [1].

Based on the questionnaire ratings, the participants were divided in their opinions of how mentally exhausting they perceived the task to be. A narrow majority in both conditions indicated a tendency towards higher ratings, implying a more tiring experience.

Table 6.2: Number of trials falsely and correctly identified as incongruent [1].

Participants	Falsely identified (of 120)	Correctly identified (of 12)
S1	0	8
S2	0	12
S3	3	9
S4	0	6
S5	0	11
S6	0	12
S7	0	11
S8	0	10
S9	0	9
S10	4	12
S11	0	12
S12	1	8
S13	0	12
S14	0	11
S15	0	8
avg	0.53	10.07
stdev	1.25	1.94

However, the task was primarily considered not physically exhausting by the majority, with none of them categorizing it as very physically tiring. There was minimal disparity between the two conditions in these aspects, with only four participants giving distinct ratings in both conditions. With respect to the ability to stay focused, the ease of concentrating on the MI task, and the vibrotactile guidance, no participant assigned the lowest rating, indicating that none found it excessively difficult to stay focused or concentrate on the task or guidance. It was observed that a higher number of participants felt it was easier to remain focused during the VtG condition, while more participants reported easier concentration on the MI task in the noVtG condition. In the VtG condition, a greater number of participants, in absolute terms, found it easy or very easy to remain focused on the task compared to the noVtG condition (60% vs. 47%, respectively). However, a combination of these results suggests that the impact of the vibrotactile guidance on these markers is largely subjective. This aligns with the casual feedback gathered from participants in earlier studies involving vibrotactile stimulation and in preliminary tests for the current study. While the results are varied in this context, Corbet et al. [56] discovered that the workload associated with an MI task, particularly frustration, effort, and

mental demand, was significantly reduced with electrotactile guidance compared to visual guidance. Additionally, Cincotti et al. [49] stated in a series of studies that compared vibrotactile feedback to visual feedback on an MI task that the majority of participants found the vibrotactile feedback more natural [1].

It is crucial to note that the behavioral data was compiled at the conclusion of the experiment, which may mean the subjective impressions of the first condition might not be entirely accurate (although the groups of participants were counterbalanced, as stated in the Chapter 3). As the behavioral results were not crucial to the primary hypotheses, these data points were gathered collectively at the end of the experiment for simplicity and brevity. For a more comprehensive analysis of behavioral data, it would be beneficial to collect data at more regular intervals (e.g., after each condition). For an extensive study of behavioral outcomes, the application of standardized questionnaires such as the System Usability Scale (SUS) [132] or the NASA Task Load Index (TLX) [133] would be more suitable [1].

6.2. Entropy Research Segment Results

The *Entropy research* segment of the dissertation presents a comparison of low-frequency amplitude and short-term entropy features derived from various TFRs of EEG MI datasets. After undergoing signal processing and feature extraction, the features were evaluated based on their sLDA classification accuracy and F1 score performance following the classification described in Chapter 5. The primary aim of this research segment was to assess the utility of short-term information entropy, obtained from a variety of TFRs, for more effective detection of MI, hence enhancing INC state detection. This outcome could potentially augment the overall performance of MI-controlled BCI systems. Moreover, the role of vibrotactile guidance in MI detection was further investigated, providing valuable insights into possible enhancements for BCI systems that incorporate somatosensory input [33].

This section thoroughly examines, analyzes, and discusses the information entropy results derived from the KGU and ULM datasets. The analysis of these datasets was conducted for different MI tasks where visual guidance alone, a mix of visual and vibrotactile guidance, or solely visual cues accompanied the MI. The evaluation of neurophysiological

patterns observed from short-term entropy and TFR results across these different MI tasks offers valuable insights into the brain's dynamics in diverse guidance contexts. The classification efforts carried out as part of this analysis included conventional low-frequency amplitude features (described in detail in Section 6.1.) as well as a novel approach using short-term Rényi and Shannon entropy calculated from six different TFRs [33].

6.2.1. Implications for Neurophysiology

The results obtained after preprocessing the EEG amplitude features for the KGU dataset, as described in Section 4.1., are presented in Fig. 6.8. The figures display the grand-average potentials for the electrode position Cz (location can be seen in Fig. 3.3), separately for each condition and each direction. Notably, slight variations can be observed in the VEPs between conditions and directions at specific time-points, including the appearance of the fixation cross ($t = -4\ s$), the presentation of the visual cue ($t = -2\ s$), and the onset of the cue motion ($t = 0\ s$). Although the observed differences in amplitude between directions are not highly prominent, they are still discernible, particularly during the MI period [33].

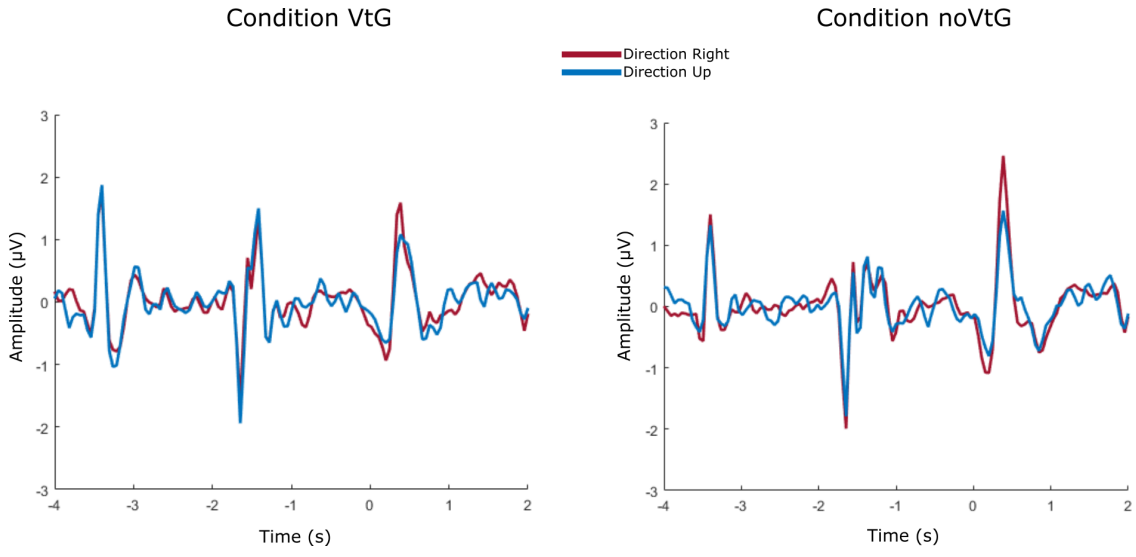


Figure 6.8: Grand-average EEG amplitude potentials for the KGU dataset, for each condition separately. Signals shown here are recorded at electrode position Cz [33].

These amplitude features served as a basis for the computation of various TFRs. Samples of outcomes for TFR, Rényi entropy, and Shannon entropy are presented in Fig. 6.9 and 6.10.

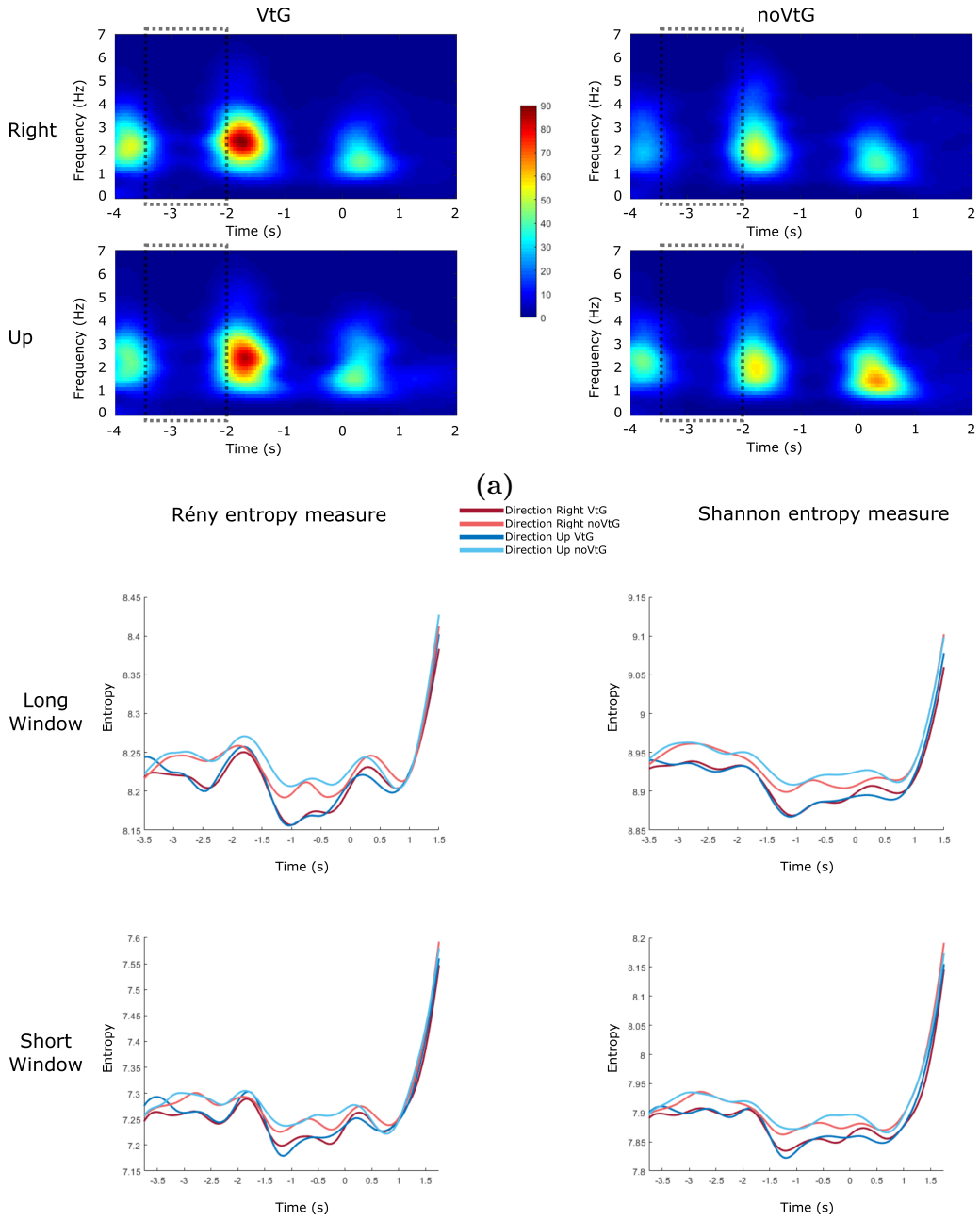
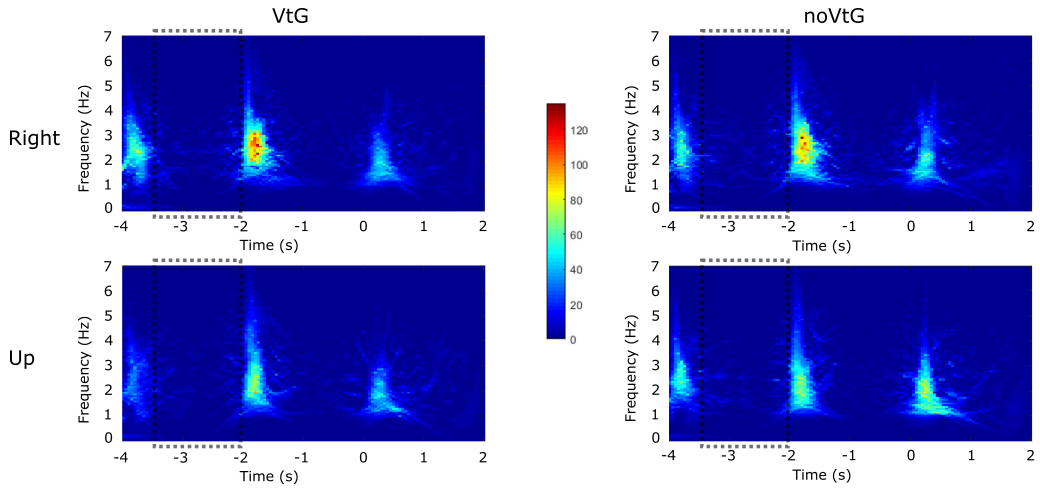
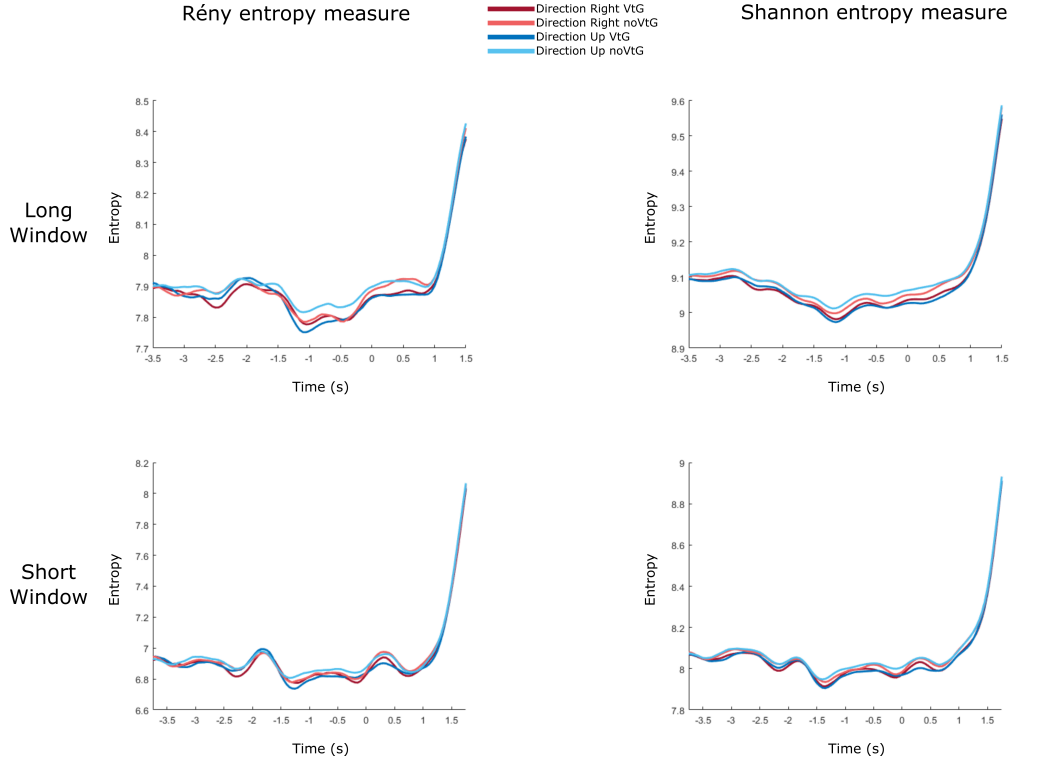


Figure 6.9: A demonstration of Spectrogram TFR, Rényi and Shannon entropy for each condition and each direction for amplitude features from the KGU dataset at the electrode position Cz. In 6.9a, a grand-average Spectrogram TFR is illustrated, and the baseline period is indicated with dashed rectangles ($t = -3.5$ s to 2 s). The grand-average results for Rényi entropy (left) and Shannon entropy (right) for the Spectrogram representation from 6.9a are shown in 6.9b for each window length (long window $w = 1$ s, and short window $w = 0.5$ s) [33].



(a)



(b)

Figure 6.10: A demonstration of Reassigned pseudo Wigner-Ville TFR, Rényi and Shannon entropy for each condition and each direction for amplitude features from the KGU dataset at the electrode position Cz. In 6.10a, a grand-average Reassigned pseudo Wigner-Ville TFR is illustrated, and the baseline period is indicated with dashed rectangles ($t = -3.5\ s$ to $-2\ s$). The grand-average results for Rényi entropy (left) and Shannon entropy (right) for the Reassigned pseudo Wigner-Ville TFR from 6.10a are shown in 6.10b for each window length (long window $w = 1\ s$, and short window $w = 0.5\ s$) [33].

The TFR employed in Fig. 6.9 is the Spectrogram. From Fig. 6.9a, it is noticeable that the VtG condition exhibits a higher magnitude in Spectrogram during the crucial time-points of the complete trial for both directions, except for the *right* direction in the noVtG condition around time $t = 0\ s$. An increase in magnitude can be seen during the initiation of the MI period ($t = 0\ s$ to $1\ s$), and a decrease towards the end of the MI period ($t = 1\ s$ to $2\ s$), which can be associated with ERS and ERD, respectively [33].

Observing the frequency range of the aforementioned changes in magnitude, these are most pronounced in the frequency band from 1 to 3.5 Hz across all conditions and directions. This finding is in line with prior research by Hehenberger et al. [1], Ofner et al. [8], and Kobler et al. [22] that posits the ERD and ERS can be decoded from low-frequency features during MI tasks.

The results for Rényi and Shannon entropy where the TFR, used as the input for entropy computation, was a Spectrogram representation are displayed in Fig. 6.9b. At equivalent time-points, Rényi and Shannon entropy produce very similar outcomes for both conditions and directions, though the entropy varies throughout the trial.

Rényi entropy is relatively low prior to and during the baseline period ($t = -3.5\ s$ to $-2\ s$), it rises at the start of the pre-MI period ($t = -2\ s$ to $-1.5\ s$), falls towards the end of the pre-MI period ($t = -1.5\ s$ to $0\ s$), and increases during the MI period ($t = 0\ s$ to $1.5\ s$). During the baseline period, Rényi entropy records lower values than during the MI period. These shifts can be associated with the change in information content at the critical time-points of the trial. This change is evident in the variation of magnitude during the trial in TFRs, which were used as an input for entropy calculation (Fig. 6.9a) [33].

For Shannon entropy, the values are at their lowest during the pre-MI period, with a less pronounced increase during the pre-MI period (compared to the baseline period).

Besides the discussed Spectrogram representation, the Reassigned pseudo Wigner-Ville, displayed in Fig. 6.10, can be inspected for comparison. It can be observed that, as with the Spectrogram representation, the VtG condition displays a stronger magnitude during critical points throughout the trial for both directions. In a manner similar to the Spectrogram representation, an increase in magnitude can be seen at the beginning of the MI period ($t = 0\ s$ to $1\ s$), followed by a decrease towards the end of the MI period ($t = 1\ s$ to $2\ s$). This pattern can be linked to ERS and ERD, as was previously identified

with the Spectrogram representation [33].

In terms of frequency, the range of these changes in magnitude is most evident in the frequency band from 1 to 4.5 Hz, for all conditions and directions [33].

The results for Rényi and Shannon entropy, where the TFR input used to compute the entropy was the Reassigned pseudo Wigner-Ville TFR, are shown in Fig. 6.10b. Mirroring the Spectrogram representation, Rényi and Shannon entropy offer very similar values for both conditions and both directions at matching time-points, but entropy varies during the trial [33].

For Rényi entropy, it is low during the baseline period ($t = -3.5\ s$ to $-2\ s$), it increases temporarily at the beginning of pre-MI ($t = -2\ s$ to $-1.5\ s$), decreases towards the end of the pre-MI period ($t = -1.5\ s$ to $0\ s$), and rises during the MI period ($t = 0\ s$ to $1.5\ s$). During the baseline period, Rényi entropy records lower values than during the MI period [33].

Likewise, for Shannon entropy, the values are lowest during the pre-MI period and peak towards the end of the MI period, albeit with fewer variations over the entire trial.

6.2.2. Feature Classification Results

Tables 6.3, 6.4, 6.5, and 6.6 present the classification results for the sLDA classifier, featuring grand-average accuracies and F1 scores during the MI period for amplitude features and Rényi entropy features with varying TFRs and window sizes. These results correspond to the long and short window datasets of KGU and ULM, respectively. Similarly, Tables 6.7, 6.8, 6.9, and 6.10 depict the same form of results for Shannon entropy features with different TFRs and window sizes for both the KGU and ULM datasets. The TFRs utilized in these tables include: Spectrogram (tfrsp), Gabor representation (tfrgabor), Pseudo Wigner-Ville (tfrpwv), Smoothed pseudo Wigner-Ville (tfrspwv), Reassigned spectrogram (tfrrsp), Reassigned Gabor spectrogram (tfrrgab), Reassigned pseudo Wigner-Ville (tfrrpwv), and Reassigned smoothed pseudo Wigner-Ville (tfrrspwv) [33].

6.2.2.1. Amplitude Features

In Table 6.3 (for the KGU dataset) and Table 6.5 (for the ULM dataset), the first column of results displays the accuracies and F1 scores for amplitude features without

entropy. The highest accuracy for the amplitude features for directions *up* versus *right* is obtained on the KGU dataset when employing amplitude features with vibrotactile guidance (condition VtG), yielding a value of 64.07%. This aligns with the earlier findings by Hehenberger et al. [1] that suggest features in VtG perform slightly better than features in noVtG (with an accuracy of 60.04%) when classifying different directions based on amplitude features. The highest overall accuracies for amplitude features, however, are reached on the KGU dataset when classifying MI (*right* or *up*) versus *base*: between 84.59% and 86.91% [33].

For the ULM dataset, the amplitude features yielded an accuracy of 53.59% when classifying different movements, which is approximately the chance level (55%). The highest accuracies for amplitude features in the ULM dataset were attained when classifying MI - EE (*elbow extension*) or EF (*elbow flexion*) - versus *base*: between 66.15% and 66.26% (shown in Table 6.5). These results are congruent with the study from which this dataset originated, which reported an accuracy of 68% (std 8%) [8].

The differential performance of the proposed methods on the amplitude features of the KGU dataset versus the ULM dataset could be attributable to several factors: the paradigms employed are not identical. They diverge in terms of timing and the types of imagined movements (movements *right* and *up* for the KGU dataset versus *elbow extension* and *elbow flexion* for the ULM dataset). Furthermore, the KGU dataset paradigm includes vibrotactile guidance on certain trials, which could potentially enhance participant engagement with the task. The superior performance of the KGU dataset could also be linked to the positive impact of visual guidance (KGU dataset) compared to the presence of only a visual cue (ULM dataset) [134]. Another possible reason for the performance disparity is the availability of different electrode positions, as discussed in 3.2. [33].

6.2.2.2. Rényi Entropy Features

As shown in Table 6.3, the KGU dataset yields the highest accuracy for long-window Rényi entropy features when utilizing the Reassigned pseudo Wigner-Ville representation TFR. This results in an accuracy of 88.29% for the vibrotactile guidance condition (VtG) *right* versus *base*. It is important to note that this represents a modest increase of 1.31% in comparison to the amplitude VtG features *right* versus *base* (which has an accuracy of

86.91%) [33].

Table 6.3: A detailed overview of sLDA classification performance using long window ($w = 1\ s$) features from the KGU dataset. Grand-average of accuracy and F1 score, categorized by feature types, and TFRs calculated with Rényi entropy. The best results per feature type are in bold [33].

cond		accuracy (%)								
TFR		-	tfrsp	tfrgabor	tfrpwv	tfrspwv	tfrrsp	tfrrgab	tfrrpwv	tfrrspwv
Feature type		ampl.	entropy	entropy	entropy	entropy	entropy	entropy	entropy	entropy
up vs. right	VtG	**64.07**	52.87	53.62	53.61	52.74	54.83	53.96	53.67	52.56
	noVtG	**60.44**	52.72	52.56	52.38	51.49	53.33	54.21	51.10	52.64
right vs. base	VtG	86.91	64.68	69.97	75.04	66.92	60.78	66.95	**88.29**	64.52
	noVtG	85.68	64.85	67.42	76.25	67.84	61.48	68.47	**87.93**	63.43
up vs. base	VtG	85.81	60.48	67.20	71.94	63.55	59.90	67.29	**88.19**	60.93
	noVtG	84.59	65.05	67.72	77.77	68.30	61.91	69.57	**88.19**	64.40
		F1 score (%)								
up vs. right	VtG	**62.83**	51.08	53.35	52.15	51.28	54.41	52.95	54.20	51.57
	noVtG	**59.04**	51.96	52.28	50.92	50.88	52.69	53.62	49.78	51.94
right vs. base	VtG	84.19	64.44	69.87	72.15	66.50	60.35	66.54	**87.30**	64.23
	noVtG	82.52	63.60	67.34	73.35	67.21	61.31	67.55	**86.69**	63.86
up vs. base	VtG	82.30	59.76	67.50	68.59	63.17	59.49	66.47	**86.93**	60.47
	noVtG	80.55	64.14	67.87	75.03	67.40	61.84	68.50	**87.06**	63.84

The highest accuracy achieved for short-window Rényi entropy features on the KGU dataset, and overall for Rényi entropy, is 98.78%. This result was achieved under the noVtG condition, specifically for the direction *right* versus *base*, when utilizing the Reassigned pseudo Wigner-Ville TFR, as demonstrated in Table 6.4. This represents a significant increase of 13.10% compared to the amplitude noVtG features *right* versus *base*, which yielded an accuracy of 85.68% [33].

Table 6.4: A detailed overview of sLDA classification performance using short window ($w = 0.5$ s) features from the KGU dataset. Grand-average of accuracy and F1 score, categorized by feature types, and TFRs calculated with Rényi entropy. The best results per feature type are in bold [33].

cond		accuracy (%)							
TFR		tfrsp	tfrgabor	tfrpwv	tfrspwv	tfrrsp	tfrrgab	tfrrpwv	tfrrspwv
Feature type		entropy	entropy	entropy	entropy	entropy	entropy	entropy	entropy
up vs. right	VtG	51.06	53.43	54.49	52.22	53.37	52.86	**55.24**	52.56
	noVtG	52.37	52.49	53.21	53.05	53.17	**53.27**	52.78	51.43
right vs. base	VtG	77.33	82.08	87.87	80.43	73.35	81.65	**98.39**	73.40
	noVtG	75.80	80.94	89.02	81.39	71.30	84.08	**98.78**	70.76
up vs. base	VtG	74.66	80.74	87.30	79.04	71.10	82.40	**98.44**	69.60
	noVtG	78.01	80.87	90.07	81.14	70.59	82.55	**98.65**	72.60
		F1 score (%)							
up vs. right	VtG	50.00	53.30	52.78	51.49	52.70	51.43	54.43	51.54
	noVtG	51.41	51.87	53.30	52.58	52.76	52.41	51.45	50.63
right vs. base	VtG	76.79	81.41	85.93	79.56	72.25	80.03	**98.29**	72.75
	noVtG	74.31	80.44	87.50	80.47	70.07	82.74	**98.71**	69.61
up vs. base	VtG	73.39	79.80	85.29	77.72	69.69	80.71	**98.33**	67.95
	noVtG	77.03	80.20	88.65	80.03	69.27	81.17	**98.55**	71.96

The highest accuracy achieved for long-window Rényi entropy features on the ULM dataset is 71.28%. This was achieved for the EE versus *base*, specifically using the Reassigned pseudo Wigner-Ville TFR, as seen in Table 6.5. This accuracy represents an increase of 5.02% compared to the results obtained using amplitude features for EE versus *base*, which yielded an accuracy of 66.26% [33].

The highest accuracy for short window Rényi entropy features on the ULM dataset, and indeed the overall highest Rényi entropy accuracy on this dataset, was 87.17%. This outcome was observed for EF versus *base* using the Reassigned pseudo Wigner-Ville TFR, as demonstrated in Table 6.6. This value reflects a substantial increase of 21.02% when compared to the accuracy achieved using amplitude features for EF versus *base* (66.15%) [33].

Table 6.5: A detailed overview of sLDA classification performance using long window ($w = 1$ s) features from the ULM dataset. Grand-average of accuracy and F1 score, categorized by feature types, and TFRs calculated with Rényi entropy. The best results per feature type are in bold [33].

					accuracy (%)				
TFR	-	tfrsp	tfrgabor	tfrpwv	tfrspwv	tfrrsp	tfrrgab	tfrrpwv	tfrrspwv
Feature type	ampl.	entropy	entropy	entropy	entropy	entropy	entropy	entropy	entropy
EE vs. EF	53.59	52.49	52.60	52.96	53.67	**54.23**	53.65	51.74	52.84
EE vs. base	66.26	55.13	54.05	58.6	56.42	57.21	59.35	**71.28**	58.31
EF vs. base	66.15	55.23	54.58	59.9	57.01	56.85	58.61	**70.95**	57.64
					F1 score (%)				
EE vs. EF	52.72	52.42	52.27	52.46	53.04	53.90	53.08	50.99	52.41
EE vs. base	55.85	54.68	52.8	56.34	54.7	56.01	58.56	**70.7**	56.1
EF vs. base	56.18	54.61	53.93	57.32	55.16	55.91	58.27	**70.39**	56.14

Table 6.6: A detailed overview of sLDA classification performance using short window ($w = 0.5$ s) features from the ULM dataset. Grand-average of accuracy and F1 score, categorized by feature types, and TFRs calculated with Rényi entropy. The best results per feature type are in bold [33].

					accuracy (%)			
TFR	tfrsp	tfrgabor	tfrpwv	tfrspwv	tfrrsp	tfrrgab	tfrrpwv	tfrrspwv
Feature type	entropy	entropy	entropy	entropy	entropy	entropy	entropy	entropy
EE vs. EF	53.37	52.51	51.42	52.91	**55.97**	54.61	53.85	53.19
EE vs. base	63.13	63	72.39	66.88	67.76	76.25	**87.01**	58.23
EF vs. base	64.2	62.55	72.24	65.87	68.68	76.43	**87.17**	58.37
					F1 score (%)			
EE vs. EF	51.81	51.78	51.74	52.18	55.42	54.83	53.31	52.33
EE vs. base	61.16	61.16	69.6	65.85	65.72	74.48	**85.67**	53.65
EF vs. base	62.41	60.45	69.7	64.47	66.87	74.29	**85.73**	53.23

6.2.2.3. Shannon Entropy Features

The best accuracies and F1 scores for both long and short window Shannon entropy features, like Rényi entropy, were achieved with the Reassigned pseudo Wigner-Ville TFR. This holds true for both the KGU and ULM datasets, as shown in Tables 6.7, 6.8, 6.9, and 6.10 [33].

Specifically for the KGU dataset, the highest accuracy achieved with long window Shannon entropy features was 94.21%. This result was achieved for the VtG condition for *right* vs. *base* using the Reassigned pseudo Wigner-Ville TFR, as shown in Table 6.7. This represents an improvement of 7.30% when compared to the accuracy obtained with amplitude VtG features for *right* vs. *base* (86.91%). Additionally, it is an increase of 5.92% when compared to the long window Rényi entropy values using the Reassigned

pseudo Wigner-Ville representation for amplitude VtG features *right* vs. *base* (88.29%) [33].

Table 6.7: A detailed overview of sLDA classification performance using long window ($w = 1$ s) features from the KGU dataset. Grand-average of accuracy and F1 score, categorized by feature types, and TFRs calculated with Shannon entropy. The best results per feature type are in bold [33].

cond		accuracy (%)							
TFR		tfrsp	tfrgabor	tfrpwv	tfrspwv	tfrrsp	tfrrgab	tfrrpwv	tfrrspwv
Feature type		entropy	entropy	entropy	entropy	entropy	entropy	entropy	entropy
up vs. right	VtG	**53.69**	53.14	52.83	52.41	52.73	51.51	52.74	53.01
	noVtG	53.38	52.42	52.71	52.88	52.78	**53.52**	53.19	51.99
right vs. base	VtG	65.22	79.40	83.24	71.41	59.51	58.68	**94.21**	62.65
	noVtG	65.40	77.82	85.32	71.52	59.29	58.38	**94.19**	63.41
up vs. base	VtG	62.54	77.30	81.64	68.39	60.97	58.04	**93.77**	61.69
	noVtG	67.60	79.25	86.23	72.41	58.41	57.83	**94.17**	61.32
		F1 score (%)							
up vs. right	VtG	52.51	52.72	52.19	51.16	51.84	50.07	52.45	52.00
	noVtG	53.16	52.05	51.78	53.23	52.92	52.82	51.85	51.42
right vs. base	VtG	64.29	78.91	81.30	70.98	58.48	57.51	**93.83**	61.84
	noVtG	63.91	77.22	83.54	71.01	58.07	57.80	**93.76**	62.81
up vs. base	VtG	62.03	76.76	79.44	67.83	59.24	57.00	**93.32**	61.04
	noVtG	66.73	78.73	84.87	71.81	57.43	57.26	**93.78**	60.76

The maximal accuracy for short window Shannon entropy features on the KGU dataset, which also represents the highest overall accuracy, is achieved by the noVtG features for the *right* vs. *base* direction. This was computed with the Reassigned pseudo Wigner-Ville TFR and reached 99.87% (as shown in Table 6.8). When compared with amplitude noVtG features for the *right* vs. *base* direction (which achieved 85.68%), this represents an increase of 14.19% [33].

When considering the ULM dataset, the maximum accuracy achieved for long window Shannon entropy features is 76.86%. This is obtained for the EE vs. *base* classification using the Reassigned pseudo Wigner-Ville TFR method, as displayed in Table 6.9. This result signifies an improvement of 10.60% compared to the amplitude features for EE vs. *base*, which achieved 66.26% [33].

When considering the ULM dataset, the Reassigned pseudo Wigner-Ville TFR method yields the highest accuracy for short window Shannon entropy features, achieving an accuracy of 95.27% for the classification of movement EF vs. *base*. This information can be seen in Table 6.10. Compared to the amplitude features for EF vs. *base* which had an accuracy of 66.15%, this result represents an increase of 29.12% [33]. Notably, this finding surpasses previous research on the same dataset, including studies by [8, 79, 80].

Table 6.8: A detailed overview of sLDA classification performance using short window ($w = 0.5$ s) features from the KGU dataset. Grand-average of accuracy and F1 score, categorized by feature types, and TFRs calculated with Shannon entropy. The best results per feature type are in bold [33].

cond		accuracy (%)							
TFR		tfrsp	tfrgabor	tfrpwv	tfrspwv	tfrrsp	tfrrgab	tfrrpwv	tfrrspwv
Feature type		entropy	entropy	entropy	entropy	entropy	entropy	entropy	entropy
up vs. right	VtG	52.40	52.92	52.82	52.54	52.17	51.93	**54.24**	53.27
	noVtG	52.27	53.12	53.08	54.13	53.09	**54.74**	53.97	52.44
right vs. base	VtG	78.62	89.09	94.37	84.16	59.26	63.56	**99.69**	68.90
	noVtG	77.54	87.72	94.88	86.86	58.79	65.29	**99.87**	67.00
up vs. base	VtG	76.33	87.58	93.86	84.41	60.01	62.62	**99.58**	64.17
	noVtG	79.74	87.56	95.14	86.96	57.54	63.18	**99.75**	67.92
		F1 score (%)							
up vs. right	VtG	50.85	52.46	51.61	51.43	51.25	51.08	53.67	52.49
	noVtG	50.69	51.70	52.25	53.58	52.52	54.36	52.76	51.57
right vs. base	VtG	77.43	88.26	93.86	83.10	55.79	62.28	**99.67**	67.52
	noVtG	75.88	86.83	94.38	86.10	55.43	64.44	**99.87**	65.59
up vs. base	VtG	74.60	86.59	93.25	83.53	56.79	61.15	**99.56**	62.43
	noVtG	78.55	86.65	94.61	85.95	53.81	62.03	**99.75**	66.65

Table 6.9: A detailed overview of sLDA classification performance using long window ($w = 1$ s) features from the ULM dataset. Grand-average of accuracy and F1 score, categorized by feature types, and TFRs calculated with Shannon entropy. The best results per feature type are in bold [33].

	accuracy (%)							
TFR	tfrsp	tfrgabor	tfrpwv	tfrspwv	tfrrsp	tfrrgab	tfrrpwv	tfrrspwv
Feature type	entropy	entropy	entropy	entropy	entropy	entropy	entropy	entropy
EE vs. EF	53.14	53.36	52.34	51.86	52.9	**53.69**	52.26	51.7
EE vs. base	55.55	56.26	63.67	60.28	57.64	59.17	**76.86**	62.1
EF vs. base	55.9	57.36	63.58	59.41	58.69	58.36	**76.84**	61.87
	F1 score (%)							
EE vs. EF	53.13	53.12	52.77	51.84	53.13	53.48	51.91	51.77
EE vs. base	54.95	55.69	61.91	58.57	56.24	57.5	**76.11**	60.73
EF vs. base	55.44	57.29	61.66	57.66	56.82	56.52	**76.02**	60.41

The Shannon entropy measure, compared to Rényi entropy, exhibits a smaller increase during the baseline period relative to the increase or variation occurring during the MI period. This can be observed in Fig. 6.9 and 6.10. This differential response likely accounts for the superior accuracies and F1 scores achieved by the Shannon entropy in comparison to the Rényi entropy [33].

Inspecting Tables 6.3, 6.4, 6.7, 6.8, 6.5, 6.6, 6.9, and 6.10, it can be observed that while some Rényi and Shannon entropy features performed exceptionally well in the primary task of MI detection (achieving up to 99.87% in MI vs. base), none were notably successful in distinguishing between different directions or movements (*up* vs. *right* or EE vs. EF).

Table 6.10: A detailed overview of sLDA classification performance using short window ($w = 0.5$ s) features from the ULM dataset. Grand-average of accuracy and F1 score, categorized by feature types, and TFRs calculated with Shannon entropy. The best results per feature type are in bold [33].

accuracy (%)								
TFR	tfrsp	tfrgabor	tfrpwv	tfrspwv	tfrrsp	tfrrgab	tfrrpwv	tfrrspwv
Feature type	entropy	entropy	entropy	entropy	entropy	entropy	entropy	entropy
EE vs. EF	52.84	52.65	52.15	51.69	**53.09**	52.85	53.05	51.79
EE vs. base	63.36	70.53	84.45	70.01	59.92	62.92	**94.82**	61.77
EF vs. base	62.66	69.49	84.3	69.02	60.2	63.62	**95.27**	61.82
F1 score (%)								
EE vs. EF	51.76	51.79	53.25	51.86	52.38	52.73	53.12	52.13
EE vs. base	62.2	69.67	83.39	69.04	56.69	61.34	**94.61**	58.03
EF vs. base	60.9	68.69	83.11	68.23	57.02	61.45	**95.04**	58.28

This suggests that the complexity and informational content of TFRs specific to different directions of the same limb may not be effectively captured by these features. This interpretation aligns with the observed similarities between directions as displayed in Fig. 6.9 and 6.10 [33].

Results from this part of the research enriched understanding of MI within different guidance contexts, illustrating brain dynamics and extending the implications of these findings to broader applications of INC state detection. Classification tasks were conducted using traditional amplitude features and an innovative approach with short-term Rényi and Shannon entropy derived from six different TFRs. These investigations provide insights into the complex patterns in MI task performance and highlight the utility of different features, paving the way for more effective neuroimaging approaches and applications.

6.3. Classification Methods Comparison Research Segment Results

In the forthcoming section, the research narrows in on the systematic comparison of diverse classification methods described in Chapter 5 utilizing the two datasets — KGU and ULM. An array of established and more recent, architecturally complex classification methods was benchmarked. This involved the implementation of specific CNNs with the objective of enhancing same-limb MI classification accuracy, pushing beyond the current

boundaries of state-of-the-art techniques. This extensive comparative analysis provided an in-depth perspective on the varying performance of these classification methods when applied to the complex domain of MI.

Furthermore, the classification outcomes were scrutinized to discern the impact of the different guidance techniques employed — visual guidance alone, a combination of visual and vibrotactile guidance, or visual cues only. Additionally, the influence of different data preprocessing steps on classification accuracy for all observed methods was meticulously evaluated.

In this segment of the dissertation, the choice was made to first evaluate the methods on the ULM dataset, which, in comparison to the KGU dataset, is less intricate and more limited in size. This strategy was underpinned by a number of rationales. First, utilizing the ULM dataset as a starting point allowed for a controlled environment to conduct preliminary tests, ensuring that the methods employed were functioning optimally before transitioning to the more complex KGU dataset. The simplicity of the ULM dataset also provided an effective means to compare and validate the results obtained subsequently with the KGU dataset. Second, the ULM dataset, with its single static visual cue as a method of instructing participants, served as a useful point of reference for understanding the effects of incorporating multiple guidance types, as seen in the KGU dataset. After validating the methods and gaining insights from the ULM dataset, these methods were then applied to the KGU dataset. The KGU dataset, with its more complex paradigm incorporating dynamic visual guidance and a combination of visual and vibrotactile guidance, required a more nuanced approach and comprehensive analyses, thus making it appropriate to explore after establishing a solid groundwork with the ULM dataset. This approach ensured a structured progression in the analysis, facilitating a more effective and efficient examination of the research [27].

6.3.1. Comparison of Classification Methods and Preprocessing Frequency Bands With the ULM Dataset

A comparative analysis of the classification methods outlined in Section 5.3. was conducted on the ULM dataset, utilizing both low-frequency features (0.2 to 5 Hz) and broad-frequency features (1 to 40 Hz). As indicated in Table 6.11, the most successful

classifier for distinguishing between the EF and EE tasks of MI on the ULM dataset was the ResNet-101, with an achieved accuracy of 72.30%. In comparison, DenseNet-169 demonstrated a commendable performance with an accuracy of 66.24%, while the remaining methods hovered around the chance level of 55% [27].

Table 6.11: Grand-average (across all participants) classification accuracy of the different methods obtained on the ULM dataset with low-frequency (0.2 to 5 Hz) and broad-frequency (1 to 40 Hz) features. The best results for a given feature type are shown in bold [27].

	accuracy (%)					
Classification method	sLDA	SVM	RF	VGG-19	ResNet-101	DenseNet-169
EF vs. EE (0.2 to 5 Hz)	53.59	53.07	53.93	57.47	**72.30**	66.24
EF vs. EE (1 to 40 Hz)	54.75	55.47	54.03	56.24	**69.82**	62.94

A comparison between low-frequency and broad-frequency features showed that ResNet-101 also delivered the best accuracy, reaching up to 69.82% for broad-frequency features. There was a comparable level of performance across all methods when using broad-frequency features in relation to low-frequency features. Notably, the three simpler classifiers (sLDA, SVM, and RF) exhibited a marginal improvement when utilizing broad-frequency features, while the performance of the three CNN classifiers was slightly reduced. However, it is important to note that all the simpler classifiers showed performance close to chance level on both frequency passbands, hence the improvement was negligible [27].

To facilitate a statistical examination of the influence of the employed method and the frequency passband used in the preprocessing stage, a two-way repeated measures (RM) ANOVA was performed using the ULM analysis results. Specifically, a 6×2 RM ANOVA was used, with the within-subject factors being the *Method* (6 instances) and *Passband* (0.2 to 5 Hz, 1 to 40 Hz). The test yielded the subsequent findings [27]:

- Statistical differences were identified in mean classification accuracy across the various methods considered: $F(5, 70) = 78.281, p < .001$. The post-hoc analysis with Bonferroni correction revealed that ResNet-101 statistically outperformed all other methods, achieving an accuracy of 71.1%. The detailed results of the post-hoc pairwise comparisons can be seen in Table 6.12.

- There was no significant influence of the *Passband* on the accuracy of classification:

$F(1, 14) = 0.884, p = 0.363 > 0.05$. This implies that the distinction in classification accuracy when preprocessing the ULM dataset with a low-frequency band (0.2 to 5 Hz) or a broad-frequency band (1 to 40 Hz) does not bear statistical significance.

– Furthermore, the interaction between the $Method * Passband$ factors was not found to be statistically significant: $F(5, 70) = 1.816, p = 0.121 > 0.05$.

Table 6.12: Pairwise comparisons of methods' classification accuracy, as achieved on the ULM dataset. Values represent differences in accuracy means (%). Statistically significant differences ($p < 0.05$) are shown in bold and marked with * [27].

Method	sLDA	SVM	RF	VGG-19	ResNet-101	DenseNet-169
sLDA		−0.1	0.2	−2.7	**−16.9***	**−10.4***
SVM	0.1		0.3	−2.6	**−16.8***	**−10.3***
RF	−0.2	−0.3		−2.9	**−17.1***	**−10.6***
VGG-19	2.7	2.6	2.9		**−14.2***	**−7.7***
ResNet-101	**16.9***	**16.8***	**17.1***	**14.2***		**6.5***
DenseNet-169	**10.4***	**10.3***	**10.6***	**7.7***	**−6.5***	

The performance exhibited by ResNet-101, which, to the best of the author's knowledge, has not been previously applied in the classification of MI EEG center-out movements for the same limb, is promising. The observed superior performance may be due to several factors. Firstly, the deep architecture of ResNet-101, comprising 101 layers, enables it to detect complex patterns in data, an attribute vital in MI EEG classification, where patterns can be nuanced and difficult to identify. The effective propagation of gradients during training, facilitated by residual connections and the flexibility of the model, likely contribute to the capability of ResNet-101 to discern such patterns. Secondly, a greater number of learnable parameters characterize ResNet-101 when compared to other tested classifiers. This allows it to learn intricate data representations and more proficiently adapt to the training set, but it also requires a larger dataset for training the network. Finally, its proven success in multiple computer vision tasks, including image classification and object detection, suggests that the architecture is adept at learning intricate visual data representations, which may be particularly pertinent in MI EEG classification where data are presented as images. When compared to previous studies, Lee et al. [92] used a channel-wise variational autoencoder-based CNN to achieve an accuracy of 60% in classifying the *elbow extension* against other same-limb MI tasks.

The frequency passband analysis corroborates previous findings [1, 22], affirming the importance of the low-frequency delta band (0.2 to 5 Hz) in MI information. It also suggests that the addition of higher frequencies when processing same-limb different MIs does not enhance classification accuracy. Indeed, the use of broad-frequency features in data preprocessing may lead to a reduction in classification accuracy, as indicated by the results presented in this section [27].

6.3.2. Comparison of Classification Methods, Guidance Types, and Preprocessing Frequency Bands With the KGU Dataset

The developed approach was subjected to further testing on the KGU dataset to validate the results derived from the ULM dataset. Table 6.13 presents descriptive statistics that were acquired from the examination of the influence of both guidance type and frequency passband on the KGU dataset, using the applied classification methods. It was observed that, for low-frequency features (0.2 to 5 Hz), ResNet-101 yielded the highest accuracy, up to 70.99%. DenseNet-169, sLDA, and SVM also demonstrated good performance, achieving accuracies up to 65.61%, 64.07%, and 64.07% respectively. VGG-19 achieved comparable but lower accuracies, up to 60.05%. RF performance was approximately at chance level, with accuracies reaching up to 56.49% [27].

Table 6.13: Grand-average (across all participants) classification accuracy of the different methods obtained on the KGU dataset with low-frequency (0.2 to 5 Hz) and broad-frequency (1 to 40 Hz) features, shown by condition. The best results for a given feature type and condition are shown in bold [27].

	cond.	accuracy (%)					
Classification method		sLDA	SVM	RF	VGG-19	ResNet-101	DenseNet-169
up vs. right (0.2 to 5 Hz)	VtG	64.07	64.07	56.49	59.29	**70.99**	65.31
	noVtG	60.44	59.64	55.87	60.05	**70.15**	65.60
up vs. right (1 to 40 Hz)	VtG	60.87	59.38	56.96	55.63	**67.93**	62.13
	noVtG	57.66	55.72	54.75	55.53	**68.59**	60.50

From a comparative perspective of the guidance types presented in Table 6.13, it can be seen that for all classifiers, with the exception of VGG-19 and DenseNet-169, the accuracy under the VtG condition exceeds that of the noVtG condition. The ResNet-101 classifier achieved the peak accuracy for both VtG and noVtG (70.99% and 70.15%,

respectively). The greatest variations between the accuracies of the different guidance types were demonstrated by sLDA and SVM, with both achieving better accuracies under VtG guidance. In the case of broad-frequency features, all methods except ResNet-101 performed more favorably under VtG guidance [27], although ResNet-101 performed very similarly on both guidance types.

Regarding broad-frequency features, ResNet-101 achieved the highest accuracy, reaching up to 68.59%. All methods performed less well with broad-frequency features than with low-frequency features, except for RF, which achieved near chance level accuracy for both feature types [27].

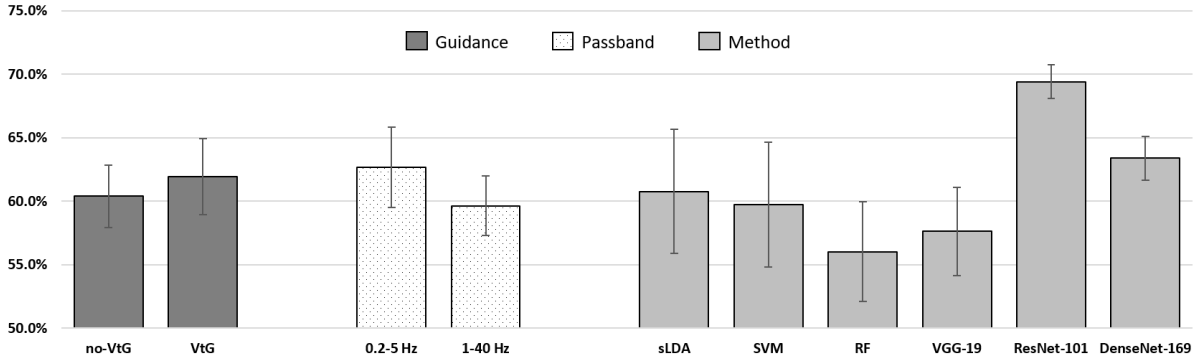


Figure 6.11: Statistical descriptions related to the classification accuracy for the KGU dataset [27].

As displayed in Fig. 6.11, the classification accuracy values taken from Table 6.13 have been averaged considering all variables under scrutiny (classification method employed, type of guidance, and passband frequency). A three-way RM ANOVA, namely a $6 \times 2 \times 2$ RM ANOVA, was performed to statistically evaluate the classification accuracy outcomes for the KGU dataset. The within-subject factors comprised *Method* (with 6 instances), *Guidance* (including VtG and noVtG), and *Passband* (0.2 to 5 Hz and 1 to 40 Hz). The Greenhouse-Geisser ε correction was applied where needed to address violations of sphericity. Whenever a significant effect was discovered, post-hoc pairwise comparisons using a Bonferroni correction were performed. The test produced the following findings [27]:

– The classification accuracy was found to be significantly influenced by the choice of *Method*: $F(2.891, 40.480) = 37.574, \varepsilon = 0.578, p < .001$. ResNet-101, similar to the results obtained for the ULM dataset, rendered the best classification accuracy

with a grand-average of 69.4%. The RF method, on the other hand, showed notably lower accuracy, with a grand-average value of 56.0%. Post-hoc pairwise comparisons with Bonferroni correction yielded results, which can be found in Table 6.14.

- *Guidance* was also revealed to have a significant impact on the classification accuracy: $F(1, 14) = 5.293, p = 0.037 < 0.05$. Therefore, the classification accuracy with vibrotactile guidance (VtG) was higher (61.9%) compared to the scenario without vibrotactile guidance (noVtG, 60.4%). Although the difference appears minimal in absolute measures, it holds statistical significance.

- Furthermore, the chosen *Passband* also had a significant influence on the classification accuracy: $F(1, 14) = 18.3, p = 0.001 < 0.05$. The accuracy was found to be significantly higher (62.7%) when the data was preprocessed using a low-frequency band compared to a broad-frequency band (59.6%).

- All interactions between the observed factors are not statistically significant:

  - $Method * Guidance$: $F(5, 70) = 1.410, p = 0.231 > 0.05, ns.$
  - $Method * Passband$: $F(5, 70) = 2.019, p = 0.086 > 0.05, ns.$
  - $Guidance * Passband$: $F(1, 14) = 0.043, p = 0.839 > 0.05, ns.$
  - $Method * Guidance * Passband$: $F(2.540, 35.553) = 0.322, \varepsilon = 0.508, p = 0.777 > 0.05, ns.$

Table 6.14: Pairwise comparisons of methods' classification accuracy, as achieved on the KGU dataset. Values represent differences in accuracy means (%). Statistically significant differences ($p < 0.05$) are shown in bold and marked with * [27].

Method	sLDA	SVM	RF	VGG-19	ResNet-101	DenseNet-169
sLDA		1.0	**4.7***	3.1	**−8.6***	−2.6
SVM	−1.0		**3.7***	2.1	**−9.7***	−3.6
RF	**−4.7***	**−3.7***		−1.6	**−13.4***	**−7.4***
VGG-19	−3.1	−2.1	1.6		**−11.8***	**−5.8***
ResNet-101	**8.6***	**9.7***	**13.4***	**11.8***		**6.0***
DenseNet-169	2.6	3.6	**7.4***	**5.8***	**−6.0***	

As previously described in Section 6.2., differences in performance when using features of the ULM dataset compared to the KGU dataset can likely be attributed to several

factors. First, the experimental paradigms used in each dataset are not identical. They differ in terms of timing (as depicted in Fig. 3.1 and Fig. 3.4) and the specific movements imagined (i.e., *elbow flexion* and *elbow extension* in the ULM dataset versus *up* and *right* in the KGU dataset). Additionally, the KGU dataset's paradigm incorporated vibrotactile guidance in certain trials, potentially leading to increased participant engagement, as discussed in Section 6.1.. The difference in performance with the KGU dataset could also be a result of the beneficial effect of visual guidance in the experiment, compared to strictly visual cues in the ULM dataset's experiment [134]. Further, differences in electrode positions (as detailed in Section 3.2.) may also contribute to the varying results [27].

Despite these differences, ResNet-101, the best-performing method applied, displayed similar performance across both datasets. This points to the robustness of the proposed preprocessing and classification pipeline. With ResNet-101, the proposed pipeline could discern different MI directions in the KGU dataset significantly better than other methods, maintaining an accuracy comparable to results obtained with the ULM dataset. These results also exceed the performance of other current CNN implementations for same-limb MI tasks, such as the 60% accuracy reported by [92], as discussed in Section 6.3.1. [27].

For the KGU dataset, the statistical analysis of different passbands indicates that using low-frequency features (0.2 to 5 Hz) in an experiment incorporating vibrotactile guidance and with the proposed preprocessing pipeline yields significantly better results than using broad-frequency features (1 to 40 Hz). This may be attributed to the fact that the low-frequency EEG amplitude modulations enhance information about arm movement initiation and directional processing as a structural part of movement execution and preparation [22] [27].

The findings from the different guidance types statistical analysis align with earlier discoveries by Hehenberger et al. [1] explained in Section 6.1., where it was demonstrated that the VtG condition does not hinder a classifier's capacity to extract directional information. On the contrary, it was found to positively impact the extraction of directional information with simple classifiers such as sLDA and SVM. In this current study, it was demonstrated that the difference between guidance types is statistically significant when the proposed preprocessing and classification pipeline is employed [27].

This comprehensive comparative study presents important results that contribute to

the ongoing discussion on the optimal methods of MI classification. The results offer new insights and encourage further research that will potentially inform and guide future advances in the field of MI studies.

Chapter 7

CONCLUSION

Brain-computer interfaces using MI as a control mode have attracted the attention of researchers around the world due to their extensive potential for real-world applications. As technological developments continue to refine sensing technology, signal processing algorithms, and machine learning, the precision and reliability of BCI systems continue to improve.

This dissertation, in its five phases, addresses three distinct but interrelated research segments, as described in the introductory chapter of this dissertation.

The initial segment of the research meticulously investigated the intricate dynamics between visual and vibrotactile guidance within the context of MI tasks. To accomplish this, a carefully designed experiment was set up, and an original dataset was curated from 15 participants. This dataset comprised of data from two same-limb center-out MI movement directions, where both vibrotactile and visual guidance were incorporated. This unique dataset provided an invaluable foundation for the subsequent investigations, offering a balanced integration of visual and tactile sensory modalities in MI-based tasks. The study demonstrated that the two same-limb center-out movement directions could be discriminated in low-frequency EEG amplitudes, with the average performance exceeding the chance level in both conditions. When vibrotactile guidance was incorporated, the average performance showed an incremental but noticeable enhancement. Remarkably, excellent average accuracies of 83.20% were reached for low-frequency amplitude features and 82.60% for mu and beta band spectral power features when classifying the MI period against the baseline. When classifying different directions, average accuracies achieved

were up to 64.07% for condition VtG (visual and vibrotactile guidance) and 60.44% for condition noVtG (visual guidance only). Even though the implementation of vibrotactile guidance did not yield a statistically significant improvement in this segment of the research, its positive impact in terms of higher and less variable average accuracies, especially for spectral power features, should not be overlooked. Hence, vibrotactile guidance emerged as a valuable adjunct, enhancing the MI extraction process without interfering with the ability to detect MI. The neurophysiological research evaluated the amplitude features and spectral power features during an average trial, under both conditions and both directions. Key findings included prominent evoked responses to the stimuli, with visual-evoked potentials present in both conditions. A somatosensory-evoked potential was observed in condition VtG, triggered by the onset of vibrotactile stimulation. An MRCP, indicative of an imagined movement initiation, presented as a central negativity after the cue movement onset was also observed. The peak amplitudes, slopes, and spatial profiles varied with movement direction and condition, with MRCP negativity being more robust and less variable in condition noVtG. Time-frequency analysis of the trials revealed a power decrease in the mu and beta frequency ranges during the MI period, with considerable inter-subject variance that might be influenced by the level of experience with MI tasks. Further, the spectral profiles showed minor variations between conditions and directions, with condition VtG exhibiting more bilateral patterns compared to the predominantly contralateral patterns in condition noVtG. Overall, the research identified subtle variations between conditions, with vibrotactile guidance seemingly influencing the spatial profile of MRCP and spectral profiles during MI. The results underline the complexity and subjectivity of the evoked responses. Furthermore, the behavioral analysis of the data gathered from the participants' questionnaire offers significant insights into participant reactions across different experiment conditions. Participants' subjective perceptions of mental and physical fatigue during the task showed mixed responses. On the physical effort spectrum, most participants reported that the task was not physically tiring. When comparing conditions with and without vibrotactile guidance, the responses indicated that the perceived effort was predominantly subjective, resulting in similar ratings across conditions. A majority of the participants indicated that they were able to maintain focus effectively throughout the task, though they found it slightly easier to concentrate on the MI task in condition noVtG. The intricate insights gained from this

segment of the research build robust groundwork for subsequent research, thereby accomplishing the third scientific contribution, which involves acquiring an original dataset of MI BCI data with kinesthetic vibrotactile guidance. This segment also sets the stage for fulfilling the first two scientific contributions of this dissertation.

In the second segment of the research, the spotlight shifted to the exploration of different feature types for the processing and classification of MI data. These encompass the conventional low-frequency amplitude features as well as an innovative approach incorporating short-term Rényi and Shannon information entropy based on six different TFRs derived from low-frequency amplitude features. In this segment, an additional ULM dataset was introduced to complement the originally curated KGU dataset. While amplitude features yielded better classification accuracy for different MI directions (with up to a grand-average of 64.07% for the KGU dataset), short-term Shannon entropy features based on Reassigned pseudo Wigner-Ville TFR were superior in distinguishing MI from the baseline period (MI detection), achieving outstanding grand-average accuracies of up to 99.87% for the KGU dataset and up to 95.27% for the ULM dataset. Importantly, grand-average accuracy in this case was enhanced by up to 14.19% and 29.12% for the KGU and ULM datasets, respectively, when compared to amplitude features. Neurophysiologically, observed low-frequency magnitude changes in various TFRs during key time points of the trial also resulted in short-term entropy measure (information content) changes, which can be connected to ERD, ERS, and MRCP, as pointed out in the previous segment. An expanded view of the intricate patterns discernible in MI task performance and an enriched understanding of its complexities are provided by the investigation of these diverse features. This underscores the potential of short-term entropy features in enhancing MI detection efficiency, an important feature for real-world BCI applications where preventing unwanted movement detection is crucial. In this segment, a second contribution has been achieved by demonstrating that the quantity of classifiable MI information, or entropy output, varies during the period of MI compared to a baseline (rest state). This finding significantly impacts the classification accuracy, showing an improvement when using entropy features over amplitude features for such purposes, thereby bolstering the ability to detect MI.

The concluding segment of the research took a deep dive into machine learning, conducting an exhaustive examination of various classifiers and their proficiency for the pur-

pose of classifying center-out imagined movements in the same-limb MI. This analysis incorporated a diverse group of six classifiers, encompassing three conventional ones as well as three cutting-edge classifiers based on CNNs, which were novel applications for this purpose. These classifiers were tested against both low-frequency and broad-frequency amplitude features, offering a well-rounded perspective on their efficacy. The ResNet-101 deep learning model, applied for the first time in this context, emerged as a standout performer, achieving significant grand-average accuracies of 72.30% for the ULM dataset and 70.99% for the KGU dataset when classifying different MI directions of low-frequency amplitude features. This was a noteworthy result that outstripped the performances of the other tested classifiers and state-of-the-art methods for this purpose. Moreover, earlier findings that the low-frequency band harbors valuable MI information and that introducing broader frequencies into the pipeline does not necessarily improve classification accuracy were fortified. Statistical analysis also led to the conclusion that the use of vibrotactile guidance could lead to significant improvements in classification. Therefore, in an experimental setting where vibrotactile guidance is deployed, significantly enhanced classification accuracy of directional information can be anticipated. The beneficial impact of this method is particularly notable when employing simpler classifiers, such as sLDA and SVM. This segment sheds light on the strengths, limitations, and unique attributes of each method, providing a clear performance benchmark. In this context, CNNs, despite their architectural complexity, demonstrated the potential to significantly improve classification accuracy beyond existing state-of-the-art methods. The closing findings from this research segment affirm the fourth contribution by demonstrating that the refined classification pipeline, when applied to both the original KGU dataset and a publicly available ULM dataset, can enhance successful identification of distinct interaction intentions in MI. Furthermore, this segment highlights the first contribution by statistically confirming the potential of vibrotactile stimulation in an MI experiment to boost the classification accuracy of interaction intentions.

Taken together, the results of these three research segments elucidate the complex dynamics within feature extraction in MI tasks and underscore the utility of novel approaches and advanced machine learning techniques in the classification of same-limb different movements in MI.

Although the advantages of vibrotactile guidance for BCI systems are obvious, the

potential limitations and challenges should also be acknowledged. Notably, individual variability in response to vibrotactile stimuli may occur due to factors such as age, sensory thresholds, and personal preferences. The effectiveness of vibrotactile stimulation may also decrease over time due to habituation after prolonged exposure, which remains to be researched. The technical challenge of incorporating a vibrotactile feedback system into a compact and user-friendly device is another potential obstacle. Adjusting the hardware to the operating position, while not overly time-consuming, can also be considered a potential limitation.

Both datasets utilized in this study were collected in an offline setting rather than in real-time BCI experiments. This study primarily examined an offline, single-session context and did not investigate potential within-session performance degradation due to the non-stationarity of EEG signals, a common phenomenon in online studies where feature shifts may occur. Although this limitation is acknowledged, future research could address this issue by using techniques such as Filter Bank Common Spatial Patterns in combination with Stationary Subspace Analysis, which are known to reduce performance issues and the dimensionality of the data, usually resulting in better classification accuracy.

The machine learning pipeline presented in this research has not been subjected to formal verification. Recently, formal methods, i.e., mathematical and logical approaches used to validate data preparation and training phases and to verify machine learning systems, have become increasingly important in the field of machine learning. Despite the use of existing datasets, well-known models and architectures, and commonly used software libraries in the study, the inclusion of a formal approach to verify AI-based classification techniques could further improve the proposed solutions.

Although the potential impact on workload and real-world applications of the findings were explored through behavioral analysis to emphasize the relevance of the research in practical scenarios, it is important to note a significant limitation: The questionnaire used was not standardized. Consequently, the user experience and real-world usability of the developed methods could not be quantitatively assessed. The potential impact of the findings and their prospective utility in practical applications, such as the development of more intuitive and user-friendly BCIs, may have been hindered by this shortcoming. To overcome this limitation, future research should consider incorporating standardized instruments, such as the System Usability Scale or the NASA Task Load Index, to fully

quantify the practical implications and usability of the methods.

These identified limitations largely shape future research plans. One focus will be to conduct an online BCI experiment to fully understand the effectiveness of the proposed solutions. It will also investigate how vibrotactile guidance affects the performance of the proposed pipeline across a variety of new activities and different MI tasks. Such an investigation could elucidate the application-specific differences in the effects of different MI tasks and guidance types on classification performance.

Finally, a well-known issue in MI studies is the quantity of data in available datasets. Experiments tend to be lengthy, and amassing a sufficient number of trials, especially for more complex classification methods, presents a significant challenge. As a future research step, the effects of data augmentation on the described datasets could be investigated.

Looking ahead, it is important that future research continue this multifaceted exploration and evaluate the applicability of these methods to a broader spectrum of MI tasks. With the goal of further refining these techniques and improvement of understanding of the complexities of BCIs, the implications of this research have far-reaching potential. By continuing on this path, even greater potential in BCI systems can be expected, expanding the horizons of this exciting field of research.

The culmination of this dissertation's findings represents a substantial advancement in comprehension of BCIs based on MI, marking a significant step towards more informed innovation in this field. The insights derived from the multifaceted exploration of vibrotactile guidance, short-term information entropy, and advanced machine learning methods serve as the bedrock for future investigations. These are crucial in understanding the complexities of MI interaction intention detection and create an enriched context for comprehending the mechanics of MI. Several novel perspectives were unveiled by the research, leading to a deeper understanding of MI characteristics and a more nuanced appreciation of the roles of different guidance types and the efficacy of various classification methodologies in this field. The findings also underscore the hypotheses of this research. The vibrotactile stimulation can enhance the classification accuracy of interaction intentions in a BCI system; the quantity of classifiable MI information (entropy output) does indeed change during MI as compared to a baseline, thus increasing classification accuracy when using entropy features; and by applying the refined classification pipeline, the detectability of MI can be further improved. Although the research surfaces

some challenges and areas requiring further exploration, it decisively paves the way for the development of more accurate, robust, and user-friendly MI detection techniques in the future. This dissertation thus offers a contribution to the field of MI research, laying the foundation for future discoveries and improvements in this exciting domain.

LITERATURE

[1] Lea Hehenberger, Luka Batistic, Andreea I Sburlea, and Gernot R Müller-Putz. Directional decoding from eeg in a center-out motor imagery task with visual and vibrotactile guidance. *Frontiers in human neuroscience*, page 548, 2021.

[2] F Pichiorri, F De Vico Fallani, F Cincotti, F Babiloni, M Molinari, SC Kleih, C Neuper, A Kübler, and D Mattia. Sensorimotor rhythm-based brain–computer interface training: the impact on motor cortical responsiveness. *Journal of neural engineering*, 8(2):025020, 2011.

[3] Floriana Pichiorri, Giovanni Morone, Manuela Petti, Jlenia Toppi, Iolanda Pisotta, Marco Molinari, Stefano Paolucci, Maurizio Inghilleri, Laura Astolfi, Febo Cincotti, et al. Brain–computer interface boosts motor imagery practice during stroke recovery. *Annals of neurology*, 77(5):851–865, 2015.

[4] Natalie Mrachacz-Kersting, Susan Aliakbaryhosseinabadi, Martin Pedersen, Ning Jiang, and Dario Farina. Tactile stimulation training to enhance mrcp detection in chronic stroke patients. In *International Conference on Augmented Cognition*, pages 354–363. Springer, 2017.

[5] Maureen Clerc, Laurent Bougrain, and Fabien Lotte. *Brain-computer interfaces 1: Methods and perspectives*. John Wiley & Sons, 2016.

[6] Dale Purves, George J Augustine, David Fitzpatrick, William Hall, Anthony-Samuel LaMantia, Leonard White, Michael L Platt, and Richard D Mooney. *Neuroscience*. Oxford university press, 2019.

[7] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis

of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[8] Patrick Ofner, Andreas Schwarz, Joana Pereira, and Gernot R Müller-Putz. Upper limb movements can be decoded from the time-domain of low-frequency eeg. *PloS one*, 12(8):e0182578, 2017.

[9] Jinbiao Liu, Yixuan Sheng, and Honghai Liu. Corticomuscular coherence and its applications: A review. *Frontiers in human neuroscience*, 13:100, 2019.

[10] Boualem Boashash. *Time-frequency signal analysis and processing: a comprehensive reference*. Academic press, 2015.

[11] Gernot R Müller-Putz, Reinmar J Kobler, Joana Pereira, Catarina Lopes-Dias, Lea Hehenberger, Valeria Mondini, Víctor Martínez-Cagigal, Nitikorn Srisrisawang, Hannah Pulferer, Luka Batistić, et al. Feel your reach: An eeg-based framework to continuously detect goal-directed movements and error processing to gate kinesthetic feedback informed artificial arm control. *Frontiers in Human Neuroscience*, page 110, 2022.

[12] Terence W Picton. The p300 wave of the human event-related potential. *Journal of clinical neurophysiology*, 9(4):456–479, 1992.

[13] Marco D Comerchero and John Polich. P3a and p3b from typical auditory and visual stimuli. *Clinical neurophysiology*, 110(1):24–30, 1999.

[14] Saša Vlahinić, Luka Batistić, Guruprasad Madhale Jadav, and Miroslav Vrankić. Brain computer interface based communicator for persons in locked-in state. *Informatica*, 30(4):781–798, 2019.

[15] Guruprasad Madhale Jadav, Luka Batistić, Saša Vlahinić, and Miroslav Vrankić. Brain computer interface communicator: a response to auditory stimuli experiment. In *2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 420–423. IEEE, 2017.

[16] Anthony M Norcia, L Gregory Appelbaum, Justin M Ales, Benoit R Cottereau,

and Bruno Rossion. The steady-state visual evoked potential in vision research: A review. *Journal of vision*, 15(6):4–4, 2015.

[17] Thomas Elbert. *Slow cortical potentials reflect the regulation of cortical excitability*. Springer, 1993.

[18] Adrien Peyrache, Nima Dehghani, Emad N Eskandar, Joseph R Madsen, William S Anderson, Jacob A Donoghue, Leigh R Hochberg, Eric Halgren, Sydney S Cash, and Alain Destexhe. Spatiotemporal dynamics of neocortical excitation and inhibition during human sleep. *Proceedings of the national academy of sciences*, 109(5):1731–1736, 2012.

[19] Bertrand Rivet, Antoine Souloumiac, Virginie Attina, and Guillaume Gibert. xdawn algorithm to enhance evoked potentials: application to brain–computer interface. *IEEE Transactions on Biomedical Engineering*, 56(8):2035–2043, 2009.

[20] Gert Pfurtscheller and FH Lopes Da Silva. Event-related eeg/meg synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110(11):1842–1857, 1999.

[21] Han Yuan and Bin He. Brain–computer interfaces using sensorimotor rhythms: current state and future perspectives. *IEEE Transactions on Biomedical Engineering*, 61(5):1425–1435, 2014.

[22] Reinmar J Kobler, Elizaveta Kolesnichenko, Andreea I Sburlea, and Gernot R Müller-Putz. Distinct cortical networks for hand movement initiation and directional processing: an eeg study. *NeuroImage*, 220:117076, 2020.

[23] Lüder Deecke, Berta Grözinger, and HH Kornhuber. Voluntary finger movement in man: Cerebral potentials and theory. *Biological cybernetics*, 23(2):99–119, 1976.

[24] Mads Jochumsen, Imran Khan Niazi, Kim Dremstrup, and Ernest Nlandu Kamavuako. Detecting and classifying three different hand movement types through electroencephalography recordings for neurorehabilitation. *Medical & biological engineering & computing*, 54(10):1491–1501, 2016.

[25] Omar Feix do Nascimento, Dario Farina, et al. Movement-related cortical potentials allow discrimination of rate of torque development in imaginary isometric plantar flexion. *IEEE Transactions on Biomedical Engineering*, 55(11):2675–2678, 2008.

[26] Han Yuan, Alexander Doud, Arvind Gururajan, and Bin He. Cortical imaging of event-related (de) synchronization during online control of brain-computer interface using minimum-norm estimates in frequency domain. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(5):425–431, 2008.

[27] Luka Batistić, Diego Sušanj, Domagoj Pinčić, and Sandi Ljubic. Motor imagery classification based on eeg sensing with visual and vibrotactile guidance. *Sensors*, 23(11), 2023. ISSN 1424-8220. doi: 10.3390/s23115064. URL `https://www.mdpi.com/1424-8220/23/11/5064`.

[28] Jonathan R Wolpaw, Dennis J McFarland, Gregory W Neat, and Catherine A Forneris. An eeg-based brain-computer interface for cursor control. *Electroencephalography and clinical neurophysiology*, 78(3):252–259, 1991.

[29] Jonathan R Wolpaw and Dennis J McFarland. Multichannel eeg-based brain-computer communication. *Electroencephalography and clinical Neurophysiology*, 90 (6):444–449, 1994.

[30] Audrey S Royer, Alexander J Doud, Minn L Rose, and Bin He. Eeg control of a virtual helicopter in 3-dimensional space using intelligent control strategies. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(6):581–589, 2010. doi: 10.1109/TNSRE.2010.2077654.

[31] Alexander J Doud, John P Lucas, Marc T Pisansky, and Bin He. Continuous three-dimensional control of a virtual helicopter using a motor imagery based brain-computer interface. *PloS one*, 6(10):e26322, 2011.

[32] Han Yuan, Tao Liu, Rebecca Szarkowski, Cristina Rios, James Ashe, and Bin He. 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. *Neuroimage*, 49(3):2596–2606, 2010.

[33] Luka Batistić, Jonatan Lerga, and Isidora Stanković. Detection of motor imagery based on short-term entropy of time–frequency representations. *BioMedical Engineering OnLine*, 22(1):1–23, 2023.

[34] Chuong H Nguyen, George K Karavas, and Panagiotis Artemiadis. Adaptive multi-degree of freedom brain computer interface using online feedback: Towards novel methods and metrics of mutual adaptation between humans and machines for bci. *PloS one*, 14(3):e0212620, 2019.

[35] Lea Hehenberger, Andreea I Sburlea, and Gernot R Müller-Putz. Assessing the impact of vibrotactile kinaesthetic feedback on electroencephalographic signals in a center-out task. *Journal of Neural Engineering*, 17(5):056032, 2020.

[36] Uwe Proske and Simon C Gandevia. The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. *Physiological reviews*, 2012.

[37] Kenneth O Johnson. The roles and functions of cutaneous mechanoreceptors. *Current opinion in neurobiology*, 11(4):455–461, 2001.

[38] Natalie Mrachacz-Kersting, Michael Voigt, Andrew James Thomas Stevenson, Susan Aliakbaryhosseinabadi, Ning Jiang, Kim Dremstrup, and Dario Farina. The effect of type of afferent feedback timed with motor imagery on the induction of cortical plasticity. *Brain research*, 1674:91–100, 2017.

[39] Roland S Johansson and Goran Westling. Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Experimental brain research*, 56(3):550–564, 1984.

[40] Ferran Galán, Mark R Baker, Kai Alter, and Stuart N Baker. Degraded eeg decoding of wrist movements in absence of kinaesthetic feedback. *Human brain mapping*, 36 (2):643–654, 2015.

[41] Elaine A Biddiss and Tom T Chau. Upper limb prosthesis use and abandonment: a survey of the last 25 years. *Prosthetics and orthotics international*, 31(3):236–257, 2007.

[42] Christian Pylatiuk, Stefan Schulz, and Leonhard Döderlein. Results of an internet survey of myoelectric prosthetic hand users. *Prosthetics and orthotics international*, 31(4):362–370, 2007.

[43] Sören Lewis, Michael Friedrich Russold, Hans Dietl, and Eugenijus Kaniusas. User demands for sensory feedback in upper extremity prostheses. In *2012 IEEE international symposium on medical measurements and applications proceedings*, pages 1–4. Ieee, 2012.

[44] Francesca Cordella, Anna Lisa Ciancio, Rinaldo Sacchetti, Angelo Davalli, Andrea Giovanni Cutti, Eugenio Guglielmelli, and Loredana Zollo. Literature review on needs of upper limb prosthesis users. *Frontiers in neuroscience*, 10:209, 2016.

[45] Andreas Schwarz, Patrick Ofner, Joana Pereira, Andreea Ioana Sburlea, and Gernot R Müller-Putz. Decoding natural reach-and-grasp actions from human eeg. *Journal of neural engineering*, 15(1):016005, 2017.

[46] Joana Pereira, Patrick Ofner, Andreas Schwarz, Andreea Ioana Sburlea, and Gernot R Müller-Putz. Eeg neural correlates of goal-directed movement intention. *Neuroimage*, 149:129–140, 2017.

[47] Joana Pereira, Andreea Ioana Sburlea, and Gernot R Müller-Putz. Eeg patterns of self-paced movement imaginations towards externally-cued and internally-selected targets. *Scientific reports*, 8(1):1–15, 2018.

[48] Andreea Ioana Sburlea, Luis Montesano, and Javier Minguez. Continuous detection of the self-initiated walking pre-movement state from eeg correlates without session-to-session recalibration. *Journal of neural engineering*, 12(3):036007, 2015.

[49] Febo Cincotti, Laura Kauhanen, Fabio Aloise, Tapio Palomäki, Nicholas Caporusso, Pasi Jylänki, Donatella Mattia, Fabio Babiloni, Gerolf Vanacker, Marnix Nuttin, et al. Vibrotactile feedback for brain-computer interface operation. *Computational intelligence and neuroscience*, 2007, 2007.

[50] Aniruddha Chatterjee, Vikram Aggarwal, Ander Ramos, Soumyadipta Acharya, and Nitish V Thakor. A brain-computer interface with vibrotactile biofeedback for haptic information. *Journal of neuroengineering and rehabilitation*, 4(1):1–12, 2007.

[51] Christian Antfolk, Christian Balkenius, Göran Lundborg, Birgitta Rosén, and Fredrik Sebelius. Design and technical construction of a tactile display for sensory feedback in a hand prosthesis system. *Biomedical engineering online*, 9(1):1–9, 2010.

[52] Robert Leeb, Kiuk Gwak, Dae-Shik Kim, et al. Freeing the visual channel by exploiting vibrotactile bci feedback. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3093–3096. IEEE, 2013.

[53] Paul Bach-y Rita and Stephen W Kercel. Sensory substitution and the human–machine interface. *Trends in cognitive sciences*, 7(12):541–546, 2003.

[54] Marta Franceschi, Lucia Seminara, Strahinja Dosen, Matija Strbac, Maurizio Valle, and Dario Farina. A system for electrotactile feedback using electronic skin and flexible matrix electrodes: experimental evaluation. *IEEE transactions on haptics*, 10(2):162–172, 2016.

[55] Febo Cincotti, Floriana Pichiorri, Pietro Aricò, Fabio Aloise, Francesco Leotta, Fabrizio de Vico Fallani, J del R Millán, Marco Molinari, and Donatella Mattia. Eeg-based brain-computer interface to support post-stroke motor rehabilitation of the upper limb. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4112–4115. IEEE, 2012.

[56] Tiffany Corbet, Iñaki Iturrate, Michael Pereira, Serafeim Perdikis, and José del R Millán. Sensory threshold neuromuscular electrical stimulation fosters motor imagery performance. *Neuroimage*, 176:268–276, 2018.

[57] Patrick E Patterson and Judd A Katz. Design and evaluation of a sensory feedback system that provides grasping pressure in a myoelectric hand. *J Rehabil Res Dev*, 29(1):1–8, 1992.

[58] Christian Antfolk, Christian Cipriani, Maria Chiara Carrozza, Christian Balkenius, Anders Björkman, Göran Lundborg, Birgitta Rosén, and Fredrik Sebelius. Transfer of tactile input from an artificial hand to the forearm: experiments in amputees

and able-bodied volunteers. *Disability and Rehabilitation: Assistive Technology*, 8 (3):249–254, 2013.

[59] Ander Ramos-Murguialday, Markus Schürholz, Vittorio Caggiano, Moritz Wildgruber, Andrea Caria, Eva Maria Hammer, Sebastian Halder, and Niels Birbaumer. Proprioceptive feedback and brain computer interface (bci) based neuroprostheses. *PloS one*, 7(10):e47048, 2012.

[60] Ander Ramos-Murguialday, Doris Broetz, Massimiliano Rea, Leonhard Läer, Özge Yilmaz, Fabricio L Brasil, Giulia Liberati, Marco R Curado, Eliana Garcia-Cossio, Alexandros Vyziotis, et al. Brain–machine interface in chronic stroke rehabilitation: a controlled study. *Annals of neurology*, 74(1):100–108, 2013.

[61] Luca Randazzo, Inaki Iturrate, Serafeim Perdikis, and J d R Millán. mano: A wearable hand exoskeleton for activities of daily living and neurorehabilitation. *IEEE Robotics and Automation Letters*, 3(1):500–507, 2017.

[62] Marc Jeannerod et al. The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain sciences*, 17(2):187–201, 1994.

[63] Aymeric Guillot, Christian Collet, Vo An Nguyen, Francine Malouin, Carol Richards, and Julien Doyon. Brain activity during visual versus kinesthetic imagery: an fmri study. *Human brain mapping*, 30(7):2157–2172, 2009.

[64] Parth Chholak, Guiomar Niso, Vladimir A Maksimenko, Semen A Kurkin, Nikita S Frolov, Elena N Pitsik, Alexander E Hramov, and Alexander N Pisarchik. Visual and kinesthetic modes affect motor imagery classification in untrained subjects. *Scientific reports*, 9(1):1–12, 2019.

[65] Gert Pfurtscheller and Christa Neuper. Motor imagery and direct brain-computer communication. *Proceedings of the IEEE*, 89(7):1123–1134, 2001.

[66] Gernot R Müller-Putz, Reinhold Scherer, Christian Brauneis, and Gert Pfurtscheller. Steady-state visual evoked potential (ssvep)-based communication: impact of harmonic frequency components. *Journal of neural engineering*, 2(4):123, 2005.

[67] Gernot R Muller-Putz, Reinhold Scherer, Christa Neuper, and Gert Pfurtscheller. Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces? *IEEE transactions on neural systems and rehabilitation engineering*, 14(1):30–37, 2006.

[68] Martin Rohm, Matthias Schneiders, Constantin Müller, Alex Kreilinger, Vera Kaiser, Gernot R Müller-Putz, and Rüdiger Rupp. Hybrid brain–computer interfaces and hybrid neuroprostheses for restoration of upper limb functions in individuals with high-level spinal cord injury. *Artificial intelligence in medicine*, 59(2): 133–142, 2013.

[69] Christa Neuper, Reinhold Scherer, Miriam Reiner, and Gert Pfurtscheller. Imagery of motor actions: Differential effects of kinesthetic and visual–motor mode of imagery in single-trial eeg. *Cognitive brain research*, 25(3):668–677, 2005.

[70] Kai Keng Ang, Cuntai Guan, Karen Sui Geok Chua, Beng Ti Ang, Christopher Kuah, Chuanchu Wang, Kok Soon Phua, Zheng Yang Chin, and Haihong Zhang. A clinical study of motor imagery-based brain-computer interface for upper limb robotic rehabilitation. In *2009 annual international conference of the IEEE engineering in medicine and biology society*, pages 5981–5984. IEEE, 2009.

[71] Maximilian Hommelsen, Matthias Schneiders, Christian Schuld, Philipp Keyl, and Rüdiger Rupp. Sensory feedback interferes with mu rhythm based detection of motor commands from electroencephalographic signals. *Frontiers in human neuroscience*, 11:523, 2017.

[72] Jerzy Neyman. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, volume 4. Univ of California Press, 1961.

[73] Richard G Baraniuk, Patrick Flandrin, Augustus JEM Janssen, and Olivier JJ Michel. Measuring time-frequency information content using the rényi entropies. *IEEE Transactions on Information theory*, 47(4):1391–1409, 2001.

[74] Selin Aviyente and William J Williams. Minimum entropy time-frequency distributions. *IEEE Signal Processing Letters*, 12(1):37–40, 2004.

[75] François Auger and Patrick Flandrin. Improving the readability of time-frequency and time-scale representations by the reassignment method. *IEEE Transactions on signal processing*, 43(5):1068–1089, 1995.

[76] Rui Zhang, Peng Xu, Rui Chen, Fali Li, Lanjin Guo, Peiyang Li, Tao Zhang, and Dezhong Yao. Predicting inter-session performance of smr-based brain–computer interface using the spectral entropy of resting-state eeg. *Brain topography*, 28(5): 680–690, 2015.

[77] L Tonin, A Cimolato, and E Menegatti. Do not move! entropy driven detection of intentional non-control during online smr-bci operations. In *Converging Clinical and Engineering Research on Neurorehabilitation II*, pages 989–993. Springer, 2017.

[78] Stefano Tortora, Gloria Beraldo, Luca Tonin, and Emanuele Menegatti. Entropy-based motion intention identification for brain-computer interface. In *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, pages 2791–2798. IEEE, 2019.

[79] Ji-Hoon Jeong, No-Sang Kwak, Cuntai Guan, and Seong-Whan Lee. Decoding movement-related cortical potentials based on subject-dependent and section-wise spectral filtering. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(3):687–698, 2020.

[80] Cosimo Ieracitano, Nadia Mammone, Amir Hussain, and Francesco Carlo Morabito. A novel explainable machine learning approach for eeg-based brain-computer interface systems. *Neural Computing and Applications*, pages 1–14, 2021.

[81] Shunfei Chen, Zhizeng Luo, and Haitao Gan. An entropy fusion method for feature extraction of eeg. *Neural Computing and Applications*, 29(10):857–863, 2018.

[82] Dipti Sawant, Vaibhavi Padwal, Jugal Joshi, Tanvi Keluskar, Ragini Lalwani, Tanushree Sharma, and Rohin Daruwala. Classification of motor imagery eeg signals using memd, csp, entropy and walsh hadamard transform. In *2019 IEEE Bombay Section Signature Conference (IBSSC)*, pages 1–6. IEEE, 2019.

[83] Na Ji, Liang Ma, Hui Dong, and Xuejun Zhang. Eeg signals feature extraction based on dwt and emd combined with approximate entropy. *Brain sciences*, 9(8): 201, 2019.

[84] David Steyrl, Reinhold Scherer, Oswin Förstner, and Gernot R Müller-Putz. Motor imagery brain-computer interfaces: random forests vs regularized lda-non-linear beats linear. In *Proceedings of the 6th international brain-computer interface conference*, pages 241–244, 2014.

[85] R Vargic, M Chlebo, and J Kacur. Human computer interaction using bci based on sensorimotor rhythm. In *2015 ieee 19th international conference on intelligent engineering systems (ines)*, pages 91–95. IEEE, 2015.

[86] Yuliang Ma, Xiaohui Ding, Qingshan She, Zhizeng Luo, Thomas Potter, and Yingchun Zhang. Classification of motor imagery eeg signals with support vector machines and particle swarm optimization. *Computational and mathematical methods in medicine*, 2016, 2016.

[87] Ranran Zhang, Xiaoyan Xiao, Zhi Liu, Wei Jiang, Jianwen Li, Yankun Cao, Jianmin Ren, Dongmei Jiang, and Lizhen Cui. A new motor imagery eeg classification method fb-trcsp+rf based on csp and random forest. *IEEE Access*, 6:44944–44950, 2018. doi: 10.1109/ACCESS.2018.2860633.

[88] Maouia Bentlemsan, ET-Tahir Zemouri, Djamel Bouchaffra, Bahia Yahya-Zoubir, and Karim Ferroudji. Random forest and filter bank common spatial patterns for eeg-based motor imagery classification. In *2014 5th International Conference on Intelligent Systems, Modelling and Simulation*, pages 235–238, 2014. doi: 10.1109/ISMS.2014.46.

[89] Zhiwen Zhang, Feng Duan, Jordi Solé-Casals, Josep Dinarès-Ferran, Andrzej Cichocki, Zhenglu Yang, and Zhe Sun. A novel deep learning approach with data augmentation to classify motor imagery signals. *IEEE Access*, 7:15945–15954, 2019. doi: 10.1109/ACCESS.2019.2895133.

[90] Yimin Hou, Shuyue Jia, Xiangmin Lun, Ziqian Hao, Yan Shi, Yang Li, Rui Zeng, and Jinglei Lv. GCNs-Net: A graph convolutional neural network approach for

decoding time-resolved eeg motor imagery signals. *IEEE Transactions on Neural Networks and Learning Systems*, 2022. doi: 10.1109/TNNLS.2022.3202569.

[91] Manfred Strahnen and Philipp Kessler. Investigation of a deep-learning based brain–computer interface with respect to a continuous control application. *IEEE Access*, 10:131090–131100, 2022.

[92] Do-Yeun Lee, Ji-Hoon Jeong, Byeong-Hoo Lee, and Seong-Whan Lee. Motor imagery classification using inter-task transfer learning via a channel-wise variational autoencoder-based convolutional neural network. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:226–237, 2022. doi: 10.1109/TNSRE.2022.3143836.

[93] Aline Roc, Léa Pillette, Jelena Mladenovic, Camille Benaroch, Bernard N'Kaoua, Camille Jeunet, and Fabien Lotte. A review of user training methods in brain computer interfaces based on mental tasks. *Journal of Neural Engineering*, 2020.

[94] Reinmar J Kobler, Andreea I Sburlea, Catarina Lopes-Dias, Andreas Schwarz, Masayuki Hirata, and Gernot R Müller-Putz. Corneo-retinal-dipole and eyelid-related eye artifacts can be corrected offline and online in electroencephalographic and magnetoencephalographic signals. *NeuroImage*, 218:117000, 2020.

[95] Swartz Center for Computational Neuroscience. Simulation and Neuroscience Application Platform, 2011. URL `https://github.com/sccn/SNAP`. Accessed 25.11.2022.

[96] Mike Goslin and Mark R Mine. The panda3d graphics engine. *Computer*, 37(10): 112–114, 2004.

[97] Ali Israr and Ivan Poupyrev. Tactile brush: drawing on skin with a tactile grid display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2019–2028, 2011.

[98] Granit Luzhnica, Sebastian Stein, Eduardo Veas, Viktoria Pammer, John Williamson, and Roderick Murray Smith. Personalising vibrotactile displays through perceptual sensitivity adjustment. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pages 66–73, 2017.

[99] Lea Hehenberger, Andreea Ioana Sburlea, and Gernot R Müller-Putz. Tuning of parameters for a vibrotactile kinaesthetic feedback system utilizing tactile illusions. In *8th Graz Brain-Computer Interface Conference 2019*, pages 244–248. Verlag der Technischen Universität Graz, 2019.

[100] GE Chatrian, E Lettich, and PL Nelson. Ten percent electrode system for topographic studies of spontaneous and evoked eeg activities. *American Journal of EEG technology*, 25(2):83–92, 1985.

[101] Scott Makeig, Anthony Bell, Tzyy-Ping Jung, and Terrence J Sejnowski. Independent component analysis of electroencephalographic data. In D Touretzky, MC Mozer, and M Hasselmo, editors, *Advances in Neural Information Processing Systems 8 (NIPS 1995)*, volume 8, pages 145–151. MIT Press, Cambridge, MA, USA, 1995.

[102] Anthony J Bell and Terrence J Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural computation*, 7(6):1129–1159, 1995.

[103] Maximilien Chaumon, Dorothy VM Bishop, and Niko A Busch. A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of neuroscience methods*, 250:47–63, 2015.

[104] Herbert Ramoser, Johannes Muller-Gerking, and Gert Pfurtscheller. Optimal spatial filtering of single trial eeg during imagined hand movement. *IEEE transactions on rehabilitation engineering*, 8(4):441–446, 2000.

[105] Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, and Klaus-Robert Muller. Optimizing spatial filters for robust eeg single-trial analysis. *IEEE Signal processing magazine*, 25(1):41–56, 2007.

[106] Kai Keng Ang, Zheng Yang Chin, Chuanchu Wang, Cuntai Guan, and Haihong Zhang. Filter bank common spatial pattern algorithm on bci competition iv datasets 2a and 2b. *Frontiers in neuroscience*, 6:39, 2012.

[107] Jean Morlet, G Arens, E Fourgeau, and D Glard. Wave propagation and sampling theory—part i: Complex signal and scattering in multilayered media. *Geophysics*, 47(2):203–221, 1982.

[108] Boualem Boashash, Nabeel Ali Khan, and Taoufik Ben-Jabeur. Time–frequency features for pattern recognition using high-resolution tfds: A tutorial review. *Digital Signal Processing*, 40:1–30, 2015.

[109] François Auger, Patrick Flandrin, Paulo Gonçalvès, and Olivier Lemoine. Time-frequency toolbox. *CNRS France-Rice University*, 46, 1996.

[110] Franz Hlawatsch, Gloria Faye Boudreaux-Bartels, et al. Linear and quadratic time-frequency signal representations. *IEEE signal processing magazine*, 9(2):21–67, 1992.

[111] Dennis Gabor. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26):429–441, 1946.

[112] AJEM Janssen. On the locus and spread of pseudo-density functions in the time-frequency plane. *Philips Journal of Research*, 37(3):79–110, 1982.

[113] François Auger and Éric Chassande-Mottin. Quadratic time-frequency analysis i: Cohen's class. *Time-Frequency Analysis: Concepts and Methods*, pages 131–163, 2008.

[114] Benjamin Blankertz, Steven Lemm, Matthias Treder, Stefan Haufe, and Klaus-Robert Müller. Single-trial analysis and classification of erp components—a tutorial. *NeuroImage*, 56(2):814–825, 2011.

[115] Stefan Haufe, Frank Meinecke, Kai Görgen, Sven Dähne, John-Dylan Haynes, Benjamin Blankertz, and Felix Bießmann. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage*, 87:96–110, 2014.

[116] Gernot R Müller-Putz, Reinhold Scherer, Clemens Brunner, Robert Leeb, and Gert Pfurtscheller. Better than random: a closer look on bci results. *International journal of bioelectromagnetism*, 10(ARTICLE):52–55, 2008.

[117] Suresh Balakrishnama and Aravind Ganapathiraju. Linear discriminant analysis-a brief tutorial. Technical Report 1998, Institute for Signal and Information Processing, 1998.

[118] Nello Cristianini and John Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, Cambridge, MA, USA, 2000. doi: 10.1017/CBO9780511801389.

[119] Aurélien Géron. *Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, Sebastopol, CA, USA, 2017. ISBN 978-1491962299.

[120] Leo Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

[121] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[122] John S Bridle. Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. In *Neurocomputing: Algorithms, architectures and applications*, pages 227–236. Springer, 1990.

[123] Yi-Tong Zhou and Rama Chellappa. Computation of optical flow using a neural network. In *ICNN*, pages 71–78, 1988.

[124] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[125] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[126] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[127] Selina C Wriessnegger, Gernot R Müller-Putz, Clemens Brunner, and Andreea I Sburlea. Inter-and intra-individual variability in brain oscillations during sports motor imagery. *Frontiers in Human Neuroscience*, 14:448, 2020.

[128] Minkyu Ahn and Sung Chan Jun. Performance variation in motor imagery brain–computer interface: a brief review. *Journal of neuroscience methods*, 243:103–110, 2015.

[129] Benjamin Blankertz, Claudia Sannelli, Sebastian Halder, Eva M Hammer, Andrea Kübler, Klaus-Robert Müller, Gabriel Curio, and Thorsten Dickhaus. Neurophysiological predictor of smr-based bci performance. *Neuroimage*, 51(4):1303–1309, 2010.

[130] Brendan Z Allison and Christa Neuper. Could anyone use a bci? In *Brain-computer interfaces*, pages 35–54. Springer, 2010.

[131] Aleksandar Miladinović, Miloš Ajčević, Joanna Jarmolowska, Uros Marusic, Marco Colussi, Giulia Silveri, Piero Paolo Battaglini, and Agostino Accardo. Effect of power feature covariance shift on bci spatial-filtering techniques: A comparative study. *Computer Methods and Programs in Biomedicine*, 198:105808, 2021.

[132] John Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[133] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[134] Cheng Yang, Lei Kong, Zhichao Zhang, Ye Tao, and Xiaoyu Chen. Exploring the visual guidance of motor imagery in sustainable brain–computer interfaces. *Sustainability*, 14(21):13844, 2022.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

The following abbreviations are used in this dissertation:

2D	Two dimensional space
3D	Three dimensional space
BCI	Brain–computer interface
CNN	Convolutional neural network
CSP	Common spatial pattern
DNN	Deep neural network
DOF	Degrees of freedom
DT	Decision tree
ECoG	Electrocorticography
EE	Elbow extension
EEG	Electroencephalography
EF	Elbow flexion
EMG	Electromiogram
EOG	Electrooculogram
ERD	Event-related desynchronization
ERP	Event-related potential
ERS	Event-related synchronization
fMRI	Functional magnetic resonance imaging
fNIRS	Functional near-infrared spectroscopy
GA	Grand-average
GAP	Global average pooling layer
GCN	Graph convolutional neural network

HCI	Human-computer interaction
IC	Independent component
ICA	Independent component analysis
IE	Information entropy
INC	Intentional non-control (state)
KGU	Kinesthetic guidance (dataset)
LDA	Linear discriminant analysis
LFP	Local field potential
MEG	Magnetoencephalography
ME	Movement execution
MI	Motor imagery
ML	Machine learning
MRCP	Movement-related cortical potential
NASA	National aeronautics and space administration
NF	Neurofeedback
noVtG	MI condition without vibrotactile stimulation
PDF	Probability density function
PSD	Power spectral density
QP	Quadratic programming
RBF	Radial basis function
ResNet	Residual Network
RF	Random forest
RM	Repeated measures
SCP	Slow cortical potentials
SEP	Somatosensory evoked potential
sLDA	LDA with shrinkage regularization
SMR	Sensorimotor rhythms
SNAP	Simulation and Neuroscience Application Platform
SNR	Signal-to-noise ratio
SSVEP	Steady-state visual evoked potentials
std	Standard deviation

STFT	Short-time Fourier transform
SUS	System usability scale
SVM	Support vector machine
TFR	Time-frequency representation
tfrgabor	Gabor TFR
tfrpwv	Pseudo Wigner-Ville TFR
tfrrgab	Reassigned Gabor TFR
tfrrpwv	Reassigned pseudo Wigner-Ville TFR
tfrrsp	Reassigned spectrogram TFR
tfrrspwv	Reassigned smoothed pseudo Wigner-Ville TFR
tfrsp	Spectrogram TFR
tfrspwv	Smoothed pseudo Wigner-Ville TFR
TLX	Task load index
ULM	Upper limb movement (dataset)
VEP	Visual evoked potential
VtG	MI condition with vibrotactile stimulation
WNN	Wavelet neural network
WVD	Wigner-Ville distribution

APPENDIX

A. KGU research questionnaire

EXPERIMENT QUESTIONNAIRE (kfb2)

- (filled out by experiment conductor) Subject code: _____ Date: _____
- Age: _____
- Sex (M/F): _____
- Left-handed or right-handed (L/R): _____
- Prior to this experiment, I participated in some sort of EEG Motor Imagery experiment.

YES NO

Following questions are related to the condition of the experiment in which **vibrotactile guidance was not present**. Please, read the questions carefully and answer each question by circling one number from 1 to 5. (1: Strongly disagree, 2: Disagree, 3: Neither agree nor disagree, 4: Agree, 5: Strongly agree)

1. Instructions about this condition of the experiment were clear and I understood them.

   1 2 3 4 5

2. Condition was mentally tiring.

   1 2 3 4 5

3. Condition was physically tiring.

   1 2 3 4 5

4. I was able to remain focused until the end of each run.

   1 2 3 4 5

5. In this condition, I found it easy to concentrate on imagining movement of my hand as the hand on the screen moved.

   1 2 3 4 5

Following questions are related to the condition of the experiment in which **vibrotactile guidance was present**. Please, read the questions carefully and answer each question by circling one number from 1 to 5. (1: Strongly disagree, 2: Disagree, 3: Neither agree nor disagree, 4: Agree, 5: Strongly agree)

1. Instructions about this condition of the experiment were clear and I understood them.

   1 2 3 4 5

2. Condition was mentally tiring.

   1 2 3 4 5

3. Condition was physically tiring.

   1 2 3 4 5

4. I was able to remain focused until the end of each run.

   1 2 3 4 5

5. In this condition, I found it easy to concentrate on imagining movement of my hand as the hand on the screen moved (while at the same time concentrating on the vibrotactile guidance).

   1 2 3 4 5

6. In this condition, I found it easy to concentrate on vibrotactile guidance and detect its direction.

   1 2 3 4 5

7. I think I detected and reported most of the mismatched trials (trials in which visual guidance was not in the same direction as vibrotactile guidance) correctly.

   1 2 3 4 5

BIOGRAPHY

Luka Batistić was born on October 21, 1992, in Rijeka, Croatia. At the Faculty of Engineering, University of Rijeka, he enrolled for undergraduate studies in Computer Science in 2011 and graduated with a bachelor's degree in 2014. He continues his education at the same faculty through graduate university studies in Computer Science, Computer Systems module, and graduated with a Master's degree in 2017.

In 2017, he was employed at the Faculty of Engineering, University of Rijeka, as an assistant in the Department of Computer Engineering. In the same year, he enrolled in postgraduate doctoral studies in Electrical Engineering and switched to postgraduate studies in Computer Science when it opened in 2018. In addition to his teaching and research duties, he also participated in the Proof of Concept project Thoughts communication for physically impaired people". He was also a CEEPUS fellow for a four-month research stay at the Institute of Neural Engineering - Graz University of Technology under the supervision of Professor Gernot R. Müller-Putz and Professor Erich Leitgeb.

LIST OF PUBLICATIONS

[1] G. Madhale Jadav, L. Batistić, S. Vlahinić and M. Vrankić, Brain computer interface communicator: A response to auditory stimuli experiment, 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2017, pp. 420-423, doi: 10.23919/MIPRO.2017.7973461.

[2] L. Batistić and M. Tomic, Overview of indoor positioning system technologies, 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2018, pp. 0473-0478, doi: 10.23919/MIPRO.2018.8400090.

[3] Vlahinić S, Batistić L, Jadav GM, Vrankić M. Brain Computer Interface Based Communicator for Persons in Locked-in State. Informatica. 2019 Jan 1;30(4):781-98.

[4] Hehenberger L, Batistic L, Sburlea AI, Müller-Putz GR. Directional decoding from EEG in a center-out motor imagery task with visual and vibrotactile guidance. Frontiers in human neuroscience. 2021:548.

[5] Müller-Putz GR, Kobler RJ, Pereira J, Lopes-Dias C, Hehenberger L, Mondini V, Martínez-Cagigal V, Srisrisawang N, Pulferer H, Batistić L, Sburlea AI. Feel your reach: An EEG-based framework to continuously detect goal-directed movements and error processing to gate kinesthetic feedback informed artificial arm control. Frontiers in Human Neuroscience. 2022:110.

[6] Batistić L, Lerga J, Stanković I. Detection of motor imagery based on short-term entropy of time–frequency representations. BioMedical Engineering OnLine. 2023 May;22(1):1-23.

[7] Batistić L, Sušanj D, Pinčić D, Ljubic S. Motor Imagery Classification Based on EEG Sensing with Visual and Vibrotactile Guidance. Sensors. 2023; 23(11):5064. https://doi.org/10.3390/s23115064