Konstantinos Barmpas

FROM THOUGHT TO ACTION

Enhancing Motor-Imagery Brain Computer Interfaces Through Deep Learning

Department of Computing
Imperial College London

IMPERIAL

From Thought to Action:

Enhancing Motor-Imagery Brain Computer Interfaces Through Deep Learning

Konstantinos Barmpas

May 2024

Supervised by
Prof. Stefanos Zafeiriou

Submitted in part fulfilment of the requirements for the degree of PhD in Computing and the Diploma of Imperial College London. This thesis is entirely my own work, and, except where otherwise indicated, describes my own research.

Department of Computing
Imperial College of Science, Technology and Medicine

Abstract

This Thesis delves into the intersection of Deep Learning and Motor-Imagery (MI) Brain-Computer Interfaces (BCIs), aiming to enhance their effectiveness and applicability through innovative approaches and frameworks.

In the first part of the Thesis, we present an extensive EEG data collection experiment, spanning two years and encompassing diverse scenarios. By meticulously designing the experimental protocols and utilizing cutting-edge EEG hardware, we captured and analyzed real and imagined movements of hundreds of participants. The first part of this Thesis also introduces a novel causal framework integrating Causal Reasoning into brainwave modeling for all BCI paradigms. Using this framework, we analyse the challenges of brainwave decoding in real-world BCI applications, exploring ways to address these challenges using general machine learning practices. This framework is the first study to combine Machine Learning and Causal Reasoning in the field of BCIs and to present a unified causal framework. Although theoretical in nature, we posit that this causal framework holds promise for future BCI enhancements and aids in pinpointing potential pitfalls within machine learning systems endeavoring to solve the complex problem of exploiting the brain.

In the second part of this Thesis, we introduce a lightweight deep neural network architecture leveraging the joint time-frequency scattering transform, resulting in improved classification performance and enhanced interpretability. Additionally, a novel subject selection framework is presented, enhancing personalised performance for subjects, especially those who initially exhibited poor performance. Finally, the issue of inter-subject variability in MI decoding, which hinders the generalization of deep models across subjects, is addressed by introducing a dynamic convolution framework, demonstrating improved generalization performance across subjects in various MI tasks.

Collectively, this Thesis presents a comprehensive exploration of key challenges in BCI research and proposes innovative solutions, both theoretical and practical, to advance the field, ultimately paving the way for more effective and robust BCIs.

Copyright Declaration

Acknowledgements

When I embarked on my journey as an undergraduate student at Imperial College London in October 2016, the idea of pursuing a PhD and crafting this Thesis seemed unimaginable. Then, in 2018, I met my supervisor, Prof. Stefanos Zafeiriou, during a 6-month placement position in his team. It was during this placement that I got properly introduced, for the first time, to the field of Artificial Intelligence. This experience gave me an understanding of the field without being thrown into the deep end straight away. After these months, I had never been more sure about the career path I wanted to follow which led me to a PhD under his supervision. I want to thank him for the opportunity he gave me, for steering my research to areas that I simply thought were far off in the future, for his trust in completing complex projects and finally, for making this journey not only rewarding but, more importantly, delightful and unforgettable. I am indebted to him for his mentorship and his friendship and I feel incredibly lucky to have met him and been his student.

I am deeply grateful to Dr. Yannis Panagakis for his guidance during my PhD journey. Without him, no chapter in this Thesis would have been possible. I sincerely thank him not only for his devoted academic support over the last 3.5 years but also for being there for me as a good friend and for his advice during academic hardships that forged a unique and strong relationship. I also want to express my gratitude and respect to Dr. Dimitrios Adamos for giving me a front-row seat to the cutting-edge developments in the field of BCIs and for repeatedly placing trust in me and my abilities. His trust and support, both academically and personally, have shaped my character and empowered me to complete my PhD studies successfully. I also extend my sincerest appreciation to Dr. Nikolaos Laskaris for introducing me to the captivating world of BCIs and for his exceptional mentorship and collaboration. His support and expertise in every chapter of this thesis have been truly invaluable. I also would like to thank all those with whom I have had the privilege of fruitful collaboration over the past few years, including Giorgos, Stelios, Siegfried and Mehdi.

This journey would not have been as fulfilling without the incredible people I met in office 351, who have now become cherished friends: Stathis, Foivos, Rolandos, Alexandros, Jiali, Michalis, and Dimitris. A special "thank you" goes to my dear friend (Dr. now) Giorgos Bouritsas, for being there for me all these years, for his advice in both academic and personal matters, for our endless conversations about politics and for all the laughs / stress that we have shared together.

Throughout these years, I have been lucky enough to have been encouraged and supported by many dear friends. Taso, Katerina, Styliana, Simon, Clem, Sofi, Emilie, Lucille, Jonas (and many others) thank you with all my heart. A special “thank you” is owed to my close friend Omar, whose unwavering encouragement has been my anchor throughout university. Without him, navigating through Imperial once (let alone twice) would have been impossible.

Many believe this PhD is mine alone but nothing could be further from the truth. “During these years, I have been encouraged, sustained, inspired and tolerated by the greatest group of friends anyone ever had” - my “second” family. Kosta, Niko, Paulo, Katerina, Anna-Maria, Sotiria, Rania and Matina, thank you for your unconditional love and support. And I would not be here if it weren’t for the most important people in my life, my whole family: my parents, my aunt, my sister, my grandparents and godparents - who have supported me throughout my entire life and have sacrificed a lot for me to pursue my dreams.

Lastly, to all readers, I hope you derive as much satisfaction from reading this Thesis as I did from writing it.

Konstantinos Barmpas

London, United Kingdom

2024

*This thesis is dedicated to my **heroes**:*

mom (Olga), dad (Giorgos) and aunt (Ntina)

"Thank you. I mean for everything. My whole life."

Contents

List of Acronyms

AI Artificial Intelligence

AR Augmented Reality

BCI Brain-Computer Interface

BMI Brain-Machine Interface

CAR Common Average Reference

CNN Convolutional Neural Network

CP Canonical Polyadic

CSP Common Spatial Patterns

CWT Continuous Wavelet Transform

DAG Directed Acyclic Graph

DL Deep Learning

EEG Electroencephalography

ERD Event-Related Desynchronization

ERP Event-Related Potential

ERS Event-Related Synchronization

FM Foundation Model

ICA Independent Component Analysis

ICM Independent Causal Mechanisms

MI Motor-Imagery

ML Machine Learning

MMI Mind-Machine Interface

SCP Slow Cortical Potential

SMR Sensorimotor Rhythms

SNR Signal-to-Noise

SSL Self-Supervised Learning

SSVEP Steady-State Visual Evoked Potentials

VR Virtual Reality

Notation

$\mathbf{x}$	1st Order (1D) Vector
$\mathbf{X}$	2nd Order (2D) Order Matrix
X	Tensor of Order ≥ 3
X	Causal Variable
$\mathcal{S}$	Distribution Shift
$x(t)$	Signal along time or 1D Time Function
$\psi_{\lambda}(t)$	Wavelet

List of Figures

List of Tables

1

Introduction

Contents

Brain

/brein/

1. - noun An organ of soft nervous tissue contained in the skull of vertebrates, functioning as the coordinating centre of sensation and intellectual and nervous activity.

2. - noun Intellectual capacity.

Source: Oxford Languages

1.1 Problem Scope and Challenges

The human brain is arguably the most complex object in the known observable universe. It is a marvel composed of countless neurons and synapses, forming the foundation of our memories, thoughts, experiences and consciousness. Together with the spinal cord, it serves as the central hub of our nervous system, governing everything from basic bodily functions (e.g. beating of the heart) to complex cognitive processes (e.g. perceiving the world around us). All these processes occur through the exchange and passing of messages between the brain and the different body organs. Over millions of years, the brain has evolved, particularly during the emergence of Homo Sapiens, to accommodate new challenges and capabilities (e.g. new dietary needs and linguistic complexity). Despite its remarkable growth, the brain's size has stabilized, while the world around us has undergone unprecedented change, primarily due to technological advancements.

The advent of technology has brought about a revolution, touching all spheres of human life, from healthcare to entertainment. Semiconductor chips have played a pivotal role, driving down the cost of personal computing and democratizing access to information and communication technologies. Today, a vast majority of the global population has access to the Internet and personal computing devices (e.g. smartphones), marking the dawn of the "Homo Digitalis" era [1].

The accessibility and affordability of personal computers have surged over time, propelled by both declining manufacturing costs and the exponential growth predicted by Moore's Law [2], where computing power doubles approximately every two years. This exponential increase in computational capacity not only entrenched personal computers in society but also laid the foundation for the emergence of Artificial Intelligence (AI), a new field in computer science.

AI, particularly in disciplines like Machine Learning (ML) and Deep Learning (DL), relies heavily on vast datasets to train algorithms, paving the way for machines to learn independently of human guidance [3]. This paradigm shift in learning methods has resulted in significant breakthroughs, from digital assistants capable of natural language processing to self-driving vehicles navigating complex environments autonomously. Notably, AI systems have surpassed human performance in various tasks, such as games like chess and Go, leading some experts to predict that AI may surpass human intelligence entirely in the future, a concept known as singularity.

Today's world is overwhelmed by more powerful computing machines and advancements in the field of AI. However, all these developments are the product of work done by human intelligence using our most powerful organ, our brain. And the question arises:

Is it possible to combine humans and machines ?

In other words, is it possible to harness the endless possibilities that AI has to offer in order to design computers that can be controlled using only our minds ? As society grasps the implications of increasingly powerful computing and AI technologies, there is growing interest in integrating human intelligence with machines. This interest has led to the emergence of the field of Brain-Computer Interface (BCI) (or Brain-Machine Interface (BMI) or Mind-Machine Interface (MMI)), which aims to facilitate direct communication between the human brain and computers. These interdisciplinary endeavours (ranging from neuroscience and medicine to engineering and computer science) promise to unlock new frontiers in human-computer interaction, forming the cornerstone of ongoing research efforts, including this doctoral Thesis.

At its core, a BCI system typically consists of hardware for neural signal acquisition (ranging from non-invasive external devices to implanted interfaces), sophisticated signal processing algorithms for decoding neural activity, and software / hardware systems for translating these decoded signals into meaningful commands or actions. Within the vast landscape of BCIs, this Thesis aims to focus on BCI systems that utilize Electroencephalography (EEG) signals and to delve into a specific paradigm, Motor-Imagery (MI) BCIs. MI-BCIs harness the brain's ability to generate neural patterns associated with imagined movements, offering a non-invasive means to translate these mental processes into actionable commands for external devices.

A crucial part of an MI-BCI is the decoding algorithm of neural activity, using either classical signal processing or machine learning techniques. Conventional machine learning techniques have faced significant issues when dealing with unprocessed raw EEG data. Researchers and neuroscientists have extensively investigated a variety of manually crafted feature types for EEG-related tasks, encompassing both temporal and frequency domains. Initial efforts were oriented around manually crafted features computed on a single-electrode basis, raising concerns about their capability to capture phenomena spanning multiple brain regions. Additionally, algorithms trained on manually crafted features often struggle with generalization across different individuals (also known as subjects), leading to the necessity of constructing only personalised models. The integration of DL techniques into the EEG data analysis has faced a slower adoption rate compared to other fields (like Computer Vision), primarily due to several challenges related to both the data itself and the subjects involved. These challenges encompass the limited availability of publicly accessible datasets, which hampers the training and validation of DL models. Additionally, factors pertaining to individual subjects, such as variations in brain anatomy and functionality, cognitive profiles, and familiarity with BCI technology, contribute to highly diverse patterns of EEG signals, even within the same individual. As a consequence, decoding MI tasks from EEG signals becomes an extremely challenging task, requiring sophisticated models to account for these challenges and achieve reliable results.

Thus, the problem of decoding MI brainwaves is twofold. On the one hand, there is a pressing need for more EEG data that closely mirrors real-world conditions beyond the confines of the laboratory, as well as a comprehensive mapping of the brainwave modeling problem through Machine Learning. Such efforts are crucial for gaining a comprehensive understanding of the intricacies and task-specific challenges involved in decoding motor-imagery brainwaves. On the other hand, advancements are required in the decoding algorithms tailored to address these task-specific complexities. This doctoral Thesis endeavours to introduce solutions that attempt to solve both problems. The first part of the Thesis concentrates on the acquisition of diverse EEG data and introduces a unified framework aiming to breakdown and analyze important challenges of brainwave modeling for all BCI paradigms. The second part of the Thesis delves into resolving specific challenges identified within MI-BCIs, aiming to contribute towards enhancing their efficacy and applicability.

1.2 Contributions and Thesis Overview

In the following section, we describe the contributions of this Thesis, in greater detail and per chapter:

- **Chapter 3. Capturing Real-World Brainwave Data:** In this chapter, we delve into our EEG data collection experiment, a groundbreaking endeavour conducted at Imperial College London. Our primary aim was to capture and analyze both real and imagined movements of users engaged in various gaming scenarios under real-world conditions. Through meticulous design choices, planning and execution, we conducted a comprehensive data collection process that involved participants interacting with diverse gaming environments, each presenting unique challenges (time-locked and self-paced cases) and stimuli. By integrating cutting-edge EEG hardware technology and various EEG headsets (ranging from research-grade to consumer-grade devices), we were able to capture data samples corresponding to both physical actions and mental simulations from a large number of participants over the span of two years.

- **Chapter 4. A Causal Perspective on Brainwave Modeling for BCIs:** In this chapter, we try to understand the problem of brainwave decoding through the lens of Causal Reasoning. Machine learning models have opened up enormous opportunities in the field of BCIs. Despite their great success, they usually face severe limitations when they are employed in real-life applications outside a controlled laboratory setting. Mixing Causal Reasoning and identifying causal relationships between variables of interest with brainwave modeling can change one's viewpoint on some of these major challenges, which can be found in various stages in the machine learning pipeline, ranging from data collection and data pre-processing to training methods and techniques. Here, we employ Causal Reasoning and present a framework aiming to breakdown and analyze important challenges of brainwave modeling for BCIs. Furthermore, we present how general machine learning practices, as well as brainwave-specific techniques, can be utilized and solve some of these identified challenges. Finally, we discuss appropriate evaluation schemes to measure these techniques' performance and efficiently compare them with other methods that will be developed in the future.

- **Chapter 5. BrainWave-Scattering Net:** In this chapter, we introduce a novel, lightweight, fully-learnable neural network architecture that relies on Gabor filters to delocalize EEG signal information into scattering decomposition paths along frequency and slow-varying temporal modulations. We utilize the proposed network in two distinct modeling settings, to build either a generic (training across subjects) or a personalised (training within a subject) classifier. In both cases, we demonstrate high performance for BrainWave-Scattering Net with considerably less trainable parameters as well as shorter training time compared to other state-of-the-art deep BCI architectures. Moreover, this network demonstrates enhanced interpretability properties emerging at the level of the temporal filtering operation and enables us to train efficient personalised BCI models with a limited amount of training data. In this chapter, we also introduce a novel subject selection framework that blends ideas from discriminative learning (based on Continuous Wavelet Transform) and graph-signal processing (over the sensor array). Through experimentation, we showcase enhanced personalised performance for MI-BCIs. Notably, it proves particularly advantageous for subjects who initially demonstrated suboptimal personalised performance.

- **Chapter 6. Improving Generalization of CNN-Based MI Decoders:** Neurophysiological processes underpinning EEG signals vary across subjects, causing covariate shifts in data distributions and hindering the generalization of deep models across subjects. In this chapter, we aim to address the challenge of inter-subject variability in MI decoding. To this end, we introduce a dynamic convolution framework to account for shifts caused by the inter-subject variability and we demonstrate improved generalization performance across subjects in various MI tasks for many well-established deep BCI architectures.

1.3 Impact and Applications

Accurate decoding of brainwaves, particularly in the paradigm of Motor-Imagery, offers a plethora of benefits across various fields, with a significant impact on human-computer interaction.

- **Rehabilitation:** For individuals undergoing rehabilitation after stroke, traumatic brain injury, or other neurological disorders, MI-BCIs serve as a new therapy tool, harnessing the brain's neuroplasticity to facilitate recovery and functional restoration. Patients can participate in targeted neurorehabilitation exercises and tasks (through repetitive motor imagery exercises guided by real-time feedback from MI-BCIs) aiming to retrain neural pathways associated with motor movement. This personalised and interactive approach not only enhances traditional rehabilitation techniques but also motivates and engages patients in their recovery journey. In addition, therapists using MI-BCIs can assess neurophysiological changes and tailor interventions accordingly. In summary, MI-BCIs for rehabilitation purposes offer an engaging and unique platform for movement training, enabling patients to regain lost motor function and achieve greater levels of independence and quality of life.

- **Individuals with Disabilities:** MI-BCIs represent a transformative technology with profound implications for individuals with disabilities through the power of thought alone, particularly those who have lost the ability to communicate and interact with the world due to severe motor impairments (for example, individuals with conditions such as spinal cord injuries, amyotrophic lateral sclerosis (ALS), or locked-in syndrome). MI-BCIs can enable users to communicate their intentions without relying on traditional motor pathways. This breakthrough technology facilitates not only communication but also allows users to engage with their environment more seamlessly. Tasks that were once daunting or unimaginable become realisable and doable, from operating digital devices and computers to controlling robotic limbs or wheelchairs. MI-BCIs not only restore a sense of independence to individuals with disabilities but also foster a more profound sense of connection and inclusion within their social networks.

- **Entertainment:** In the gaming industry, MI-BCIs introduce a new way of user interaction by allowing users to control characters and environments using their thoughts alone. This new immersive form of interaction opens many doors to innovative gameplay mechanics. Beyond gaming, MI-BCIs offer exciting opportunities in Augmented Reality (AR), Virtual Reality (VR) and Metaverse experiences, where users can navigate and interact with digital environments and items through mental commands (e.g. walk in the visual environment without being constrained by the physical space), creating truly immersive and engaging experiences. Additionally, MI-BCIs can potentially transform how we interact with music, media and other forms of entertainment content by enabling hands-free control of multimedia devices such as televisions, music players, and streaming platforms. This seamless integration of MI-BCIs into entertainment devices not only enhances accessibility for individuals with disabilities but also introduces a new dimension of interaction for healthy users as well. Finally, MI-BCIs hold promise in creative applications such as digital music composition or digital art, where users can express themselves and create artistic works using only their thoughts.

- **Sports / Competitions:** In competitive sports, MI-BCIs provide a cutting-edge tool for athletes to fine-tune their mental focus, visualization, and motor skills, ultimately leading to improved performance in the field. By incorporating MI-BCIs into their training programs, athletes can engage in targeted mental rehearsal and visualization exercises, enhancing their cognitive abilities and the execution of their physical movements. Moreover, MI-BCIs offer real-time feedback based on the athlete's brain activity, allowing coaches to tailor personalised training programs, maximizing the efficacy of the training sessions and enabling peak performance during games.

Overall, accurate Motor-Imagery brainwave decoding holds the promise of revolutionizing healthcare, technology, and augment human potential. Moreover, the work described in this Thesis can be applied as is or with minor modifications to a range of other BCI fields, including: Epilepsy Detection, Lie Detection, Emotion Recognition and many more.

Societal Impact: The decoding of brainwaves, made possible through BCI systems, presents a plethora of both promising opportunities and potential challenges. The rapid advancements in hardware technology and deep learning BCI systems underscore the urgent need for thoughtful consideration. The research outlined in this Thesis endeavours to advance BCI technology by improving brainwave decoding models to be more universally applicable, thereby amplifying the predominantly positive aspects of BCIs. However, BCIs also raise significant ethical, privacy, and societal considerations amidst their transformative potential. Issues surrounding data security, privacy, and the potential misuse of brain data demand careful consideration to ensure responsible development and deployment of BCI systems. In the field of neuroscience, the discussion around the safety and ethical implications of advances in neurotechnology and BCI technology traces back to the 2000s [4]. Since then, a number of institutions, bodies and countries have produced reports mapping the landscape of these issues (e.g. [5], [6]) and passed bills for protecting brain rights [7]. Within this Thesis, we introduce a module designed to identify individuals based on their brain activity accurately. Like fingerprint, facial recognition, or voice recognition systems, these biometric algorithms carry inherent risks and must be utilized cautiously, serving purposes that unequivocally benefit humanity.

1.4 Publications

The primary content of this Thesis stems from the author's published works in peer-reviewed journals and conferences during the author's work as a doctoral student at Imperial College London. Each chapter is structured around these publications, maintaining their original text as closely as possible to aid reader comprehension while ensuring a cohesive flow throughout the Thesis. In this section, these works are outlined and can be categorized into a) those directly relevant to the Thesis chapters and b) those where the author's involvement is acknowledged in the publication but not discussed in this Thesis.

1.4.1 Relevant Publications

The text of each chapter in this Thesis is based on the works described below. In summary, the following chapters contain:

- **Chapter 3:**

  The data collection experiment described in this chapter, titled: ***"MyBrainCommands:** Mind Reading for control"*, took place at the Department of Computing at Imperial College London. All the participants gave written consent. The experiment received approval by the Head of Department and the Research Governance and Integrity Team (SETREC number 22IC7505), 2022-2023.

- **Chapter 4:**
  - **Konstantinos Barmpas**, Yannis Panagakis, Georgios Zoumpourlis, Dimitrios A Adamos, Nikolaos Laskaris, Stefanos Zafeiriou. *"A Causal Perspective on Brainwave Modeling for BCIs".* Journal of Neural Engineering, 2024 - [8]
- **Chapter 5:**
  - **Konstantinos Barmpas**, Yannis Panagakis, Dimitrios A Adamos, Nikolaos Laskaris, Stefanos Zafeiriou. *"BrainWave-Scattering Net: a lightweight network for EEG-based motor imagery recognition".* Journal of Neural Engineering, 2023 - [9]
  - **Konstantinos Barmpas**, Yannis Panagakis, Dimitrios A Adamos, Nikolaos Laskaris, Stefanos Zafeiriou. *"Subject Selection Framework to Improve Personalised Models for Motor-Imagery BCIs via Wavelets and Graph Diffusion".* ICLR 2024 Workshop on Learning from Time Series For Health, 2024 - [10]
- **Chapter 6:**
  - **Konstantinos Barmpas**, Yannis Panagakis, Stylianos Bakas, Dimitrios A Adamos, Nikolaos Laskaris, Stefanos Zafeiriou. *"Improving generalization of CNN-based motor-imagery EEG decoders via dynamic convolutions".* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2023 - [11]
  - **Konstantinos Barmpas**, Yannis Panagakis, Dimitrios A Adamos, Nikolaos Laskaris, Stefanos Zafeiriou. *"A causal viewpoint on motor-imagery brainwave decoding".* ICLR2022 Workshop on the Elements of Reasoning: Objects, Structure and Causality, 2022 - [12]

Finally, one **patent** application was published and **granted**:

- **Konstantinos Barmpas**, Yannis Panagakis, Dimitrios A Adamos, Nikolaos Laskaris, Stefanos Zafeiriou. *"Classification of brain activity signals".*
  UK Patent Office / GB2605270 / 2024 - [13]
  European Patent Office / EP4452055 / 2024 - [14]
  WPO / WO2023148471 / 2023 - [15]

1.4.2 Other publications

In addition, during these studies, the author participated and co-authored the following works:

- Julian Suk, Lorenzo Giusti, Tamir Hemo, Miguel Lopez, **Konstantinos Barmpas**, Cristian Bodnar. *"Surfing on the neural sheaf"*. NeurIPS 2022 Workshop on Symmetry and Geometry in Neural Representations, 2022 - [16]
- Xiaoxi Wei, Aldo Faisal, Moritz Grosse-Wentrup, Alexandre Gramfort, Sylvain Chevallier, Vinay Jayaram, Camille Jeunet, Stylianos Bakas, Siegfried Ludwig, **Konstantinos Barmpas**, Mehdi Bahri, Yannis Panagakis, Nikolaos Laskaris, Dimitrios A Adamos, Stefanos Zafeiriou, William C Duong, Stephen M Gordon, Vernon J Lawhern, Maciej Śliwowski, Vincent Rouanne, Piotr Tempczyk. *"2021 BEETL Competition: Advancing Transfer Learning for Subject Independence & Heterogenous EEG Data Sets."* Proceedings of Machine Learning Research (PMLR) - NeurIPS 2021 Competitions and Demonstrations Track, 2022 - [17]
- Stylianos Bakas, Siegfried Ludwig, **Konstantinos Barmpas**, Mehdi Bahri, Yannis Panagakis, Nikolaos Laskaris, Dimitrios A Adamos, Stefanos Zafeiriou. *"Team Cogitat at NeurIPS 2021: Benchmarks for EEG Transfer Learning Competition"*. NeurIPS 2021 BEETL Competition: Benchmarks for EEG Transfer Learning - ***Winning Team***, 2021 - [18]

Finally, parts of the work described in this Thesis have been presented at the following **talks** and **outreach events**:

1. BCI Society Next Gen: Trainee Spotlight 2 Event 2022 (*Invited Speaker*)
2. Imperial Computing Research Showcase Day 2022 (*Invited Speaker*)
3. Imperial Computing Drop-In Demonstrations at The Great Exhibition Road Festival 2022
4. Science Talk at St. Paul's Girl School 2023 (*Invited Speaker*)
5. Imperial Medicine Coding The Future of Healthcare Conference 2024 (*Invited Speaker*)
6. Microsoft Student Ambassador AI Tech Talk 2024 (*Invited Speaker*)
7. Research Interview by Zak Jost at Welcome AI Overloads 2024 (*Invited Speaker*)
8. AI Panel at UCL x KCL TechSummit 2024 (*Invited Speaker*)

2

Related Work

Contents

In this chapter, we present the background theory underlying the work in this doctoral Thesis. We begin by briefly introducing the history of Brain-Computer Interfaces (BCIs) and their different stages. We then describe the different techniques that allow brain activity-related signal acquisition with emphasis given to Electroencephalogram (EEG) signals. We outline some of the important features characterizing these brain activity signals. We conclude by discussing the paradigm of Motor-Imagery (MI) and some MI decoding techniques (both classical signal processing-based and deep learning-based).

2.1 Electroencephalography (EEG) Signal

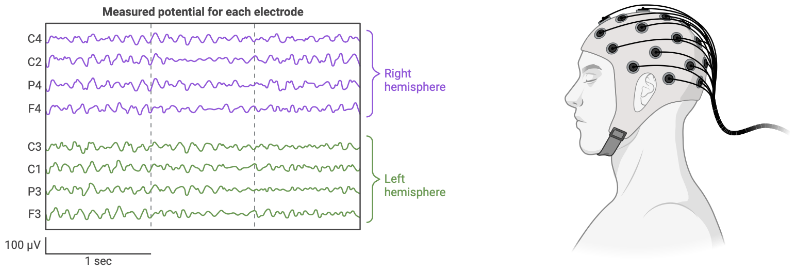

Figure 2.1: Illustration of an EEG signal.

In the 1920s, Hans Berger ([19], [20]) made a groundbreaking discovery that revolutionized our understanding of the human brain. He demonstrated that the brain generates electrical activity, a finding that laid the foundation for a new era in neuroscience. Specifically, this work illuminated how the intricate operations of neurons produce measurable electrical currents, which can be observed through various techniques, either invasive methods (small-scaled chips surgically implanted on the brain's tissue) or non-invasive techniques (electrodes placed on the scalp).

Berger's discovery resulted in a new electrophysiological monitoring technique of brain's activity: the Electroencephalogram (EEG). EEG involves placing electrodes on the scalp to detect and record the brain's electrical signals. This non-invasive method quickly became the cornerstone of brain activity monitoring, remaining the primary approach even in modern neuroscience.

EEG captures voltage fluctuations resulting from the ionic currents within the brain's neurons [21]. EEG, as demonstrated by subsequent research, has enabled neuroscientists to delve deep into understanding cognitive functions and diagnosing various neuropathologies. One of the key advantages of EEG is its non-invasive nature, making it safe and accessible for studying brain activity in diverse populations, including children and patients with neurological disorders. EEG offers a window into the brain's dynamic activity patterns, revealing rhythms and oscillations associated with different cognitive processes.

As a result, EEG is invaluable in clinical settings for diagnosing and monitoring neurological conditions such as seizures and epilepsy [22], sleep disorders [23], and brain injuries [24]. Abnormal patterns in EEG recordings can provide crucial diagnostic information, guide treatment decisions and evaluate the effectiveness of interventions.

In recent years, advancements in EEG technology, such as high-density electrode arrays and sophisticated signal processing techniques, have enhanced its utility even further. These innovations enable researchers to capture finer details of brain activity and explore complex neural networks underlying cognitive functions.

2.2 Brain-Computer Interfaces (BCIs)

Joseph Kamiya's research [25] in 1968 further expanded the potential of EEG by uncovering that specific EEG features, termed Alpha Waves, could be intentionally manipulated and controlled by human subjects. This discovery opened doors to the field of neurofeedback training, where individuals can learn to regulate their brain activity through real-time feedback. Through structured training phases, individuals can gain control over their Alpha Waves, offering promising avenues for enhancing cognitive function and addressing neurological disorders.

A few years later, Jacques Vidal suggested in his paper "Toward Direct Brain-Computer Communication" [26] that EEG signals could be used as a communication tool between humans and machines. It was in this paper that the term Brain-Computer Interface (BCI) (or BMI or MMI) was first introduced. Table 2.1 provides a summary of all significant discoveries and events in the history of Brain-Computer Interfaces (BCIs).

Year	Historic Event
1929	First EEG - (Hans Berger) [19]
1968	Alpha Waves - (Joseph Kamiya) [25]
1973	Coined the Term BCI - (Jacques Vidal) [26]
1988	P300 Speller - (Farwell and Donchin) [27]
1989	Brainwave Drawing Game - (Sobell and Trivich) [28]
1991	Sensorimotor Rhythms (SMR)-based BCI - (Wolpaw et. al.) [29]
1993	Motor Imagery (MI)-based BCI - (Pfurtscheller et. al.) [30]
1995	Steady-State Visual Evoked Potentials (SSVEP)-based BCI
1999	Slow Cortical Potential (SCP)-based BCI
2000	Common Spatial Patterns (CSP) - (Ramoser et. al.) [31]
2001	Invasive BCIs - (Nicolelis)
2007	Functional Magnetic Resonance Imaging (fMRI) - (Sitaram et. al.)
2008	Electrocorticography (ECoG) - (Schalk et. al.) [32]
2014	First Issue in the Journal "Brain-Computer Interfaces" [1]
2015	The International BCI Society was founded [2]

Table 2.1: Historical events in Brain-Computer Interface (BCI) technology. Adapted from [33].

[1]Website: https://www.tandfonline.com/journals/tbci20
[2]Website: https://bcisociety.org/

A Brain-Computer Interface is a device, or more broadly, a digital system, which enables direct communication with the external world by circumventing the usual communication pathways such as peripheral nerves or muscles [33].

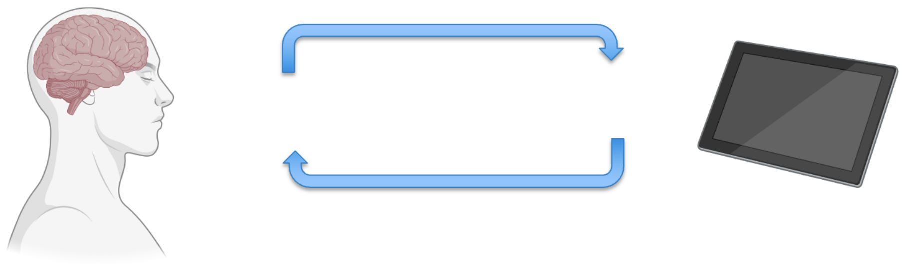

Figure 2.2: Illustration of the Brain-Computer Interface (BCI) operation loop.

The BCI operation loop consists of four main stages:

1. **Brain Activity Pattern Generation.** The necessary brain activity is generated by the human subject for the intended task.

2. **Signal Acquisition.** Brain activity signals are gathered through various methods, with the most commonly utilized ones including Electroencephalography (EEG), Functional Near-Infrared Spectroscopy (fNIRS), Functional Magnetic Resonance Imaging (fMRI), Intracortical Neuron Recording (INR), Electrocorticography (ECoG).

3. **Signal Processing.** This stage consists of three sub-stages: Signal Pre-Processing, Feature Extraction and Classification. This Thesis is focused on this stage of the BCI operation loop.

4. **Feedback back to the subject.** The Brain-Computer Interface (BCI) operation loop closes by providing feedback back to the user with the classified action by the BCI system.

2.3 Signal Acquisition and Pre-Processing

In this section, we outline several methods employed to capture brain activity signals, with a primary focus on Electroencephalogram (EEG). We also delve into the properties and patterns inherent in the EEG signals, which are crucial for developing BCIs.

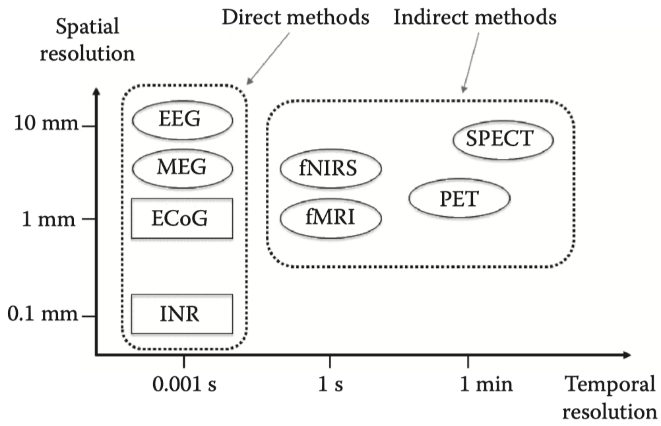

Figure 2.3: Graphical summary illustration of brain activity acquisition methods based on direct/indirect data recording and spatial / temporal resolution - Figure taken by [33].

2.3.1 Recording Methods

Brain-Computer Interfaces require a digital device that collects the brain activity signals of the subject and transfers it to a computer for processing. There are various methods that are used in the field of BCI. These techniques can be categorized based on their spatial and temporal resolution, whether they are direct or indirect as well as their invasiveness to the subject. Figure 2.3 provides a graphical summary of brain activity recording methods based on these categorization factors.

There are several recording methods (both invasive and non-invasive) of neural activity. Some of the most commonly used techniques include:

- **Functional Near-Infrared Spectroscopy (fNIRS)** [34]: It detects brain function by gauging the absorption rates of oxygenated, and deoxygenated blood, where oxygenated blood absorbs light below 800nm wavelength and deoxygenated blood absorbs light above 800nm wavelength.
- **Functional Magnetic Resonance Imaging (fMRI)** [35]: It assesses brain activity by detecting the hemodynamic response, also known as the Blood Oxygen Level–Dependent (BOLD) contrast.

Throughout this doctoral Thesis, **Electroencephalography (EEG)** [19] signals are predominantly employed. This recording technique is extensively utilized due to its non-invasive nature, cost-effectiveness, portability, and prevalence in over 60% of published research articles, as indicated by [36]. EEG measures brain's electrical activity using non-invasive sensors crafted from either gold or silver positioned on the subject's scalp. In some instances, conductive gel is applied between the EEG sensors and the scalp to heighten sensor sensitivity and facilitate more accurate detection of the subject's electrical brain activity.

The underlying principle of EEG is as follows: Every time a neuron enters the excitation state, an electrical impulse is sent from the axon to the synapse [33]. This synaptic activity generates an electrical impulse known as a postsynaptic potential. Since EEG recording takes place on the scalp, the detection of a single neuron's blast is almost impossible. But when performing a task, millions or even billions of neurons get activated synchronously which generates a electrical field, strong enough to be detected from sensors placed on the subject's scalp.

These neural voltage oscillations, which are detected by EEG sensors, are referred to as brainwaves, and based on their frequency and amplitude can be grouped into bands, as shown in Figure 2.4(b). Each brainwave band is related to a specific mental state as well as to a certain level of neurological / brain activity [33].

In order to reliably maintain the same placement of the EEG electrodes from subject to subject, the international 10/20 system [37] has been proposed for the positioning of electrodes, and it has been adopted massively since then. The 10/20 system is based on three landmarks which determine the positioning of all EEG sensors:

1. The Nasion: the connective point between the forehead and the nose
2. The Inion: the lowest point of the skull on the back of the head
3. The Pre-Auricular Points: points at the front of the ears

Under the 10/20 system, the scalp is divided into areas from the nasion to the Inion with interval rates of 10%, 20%, 20%, 20%, 20%, and 10% hence the name. These areas are named as: Frontopolar (FP), Frontal (F), Central (C), Parietal (P) and Occipital (O) respectively [37]. As it is shown in Figure 2.4(a), each site is characterized by a letter to identify the lobe (FP, F, C, P, O) followed by a number or the letter Z (even numbers refer to the right hemisphere, odd numbers refer to the left hemisphere and Z refers to the midline).

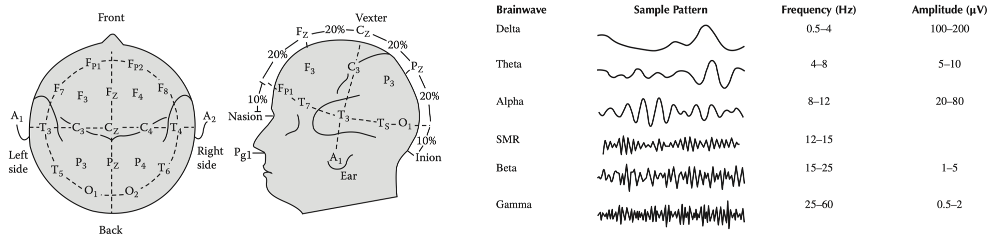

Figure 2.4: (a) The 10/20 international system of electrode placement (b) Brainwaves and frequency bands - Figures taken from [33].

2.3.2 EEG Signal Patterns

Nowadays, mainly neuroelectric signals, such as EEG, are widely used for operating BCIs. There are two main categories of EEG signals:

1. **Evoked Potential:** Evoked responses refer to EEG signals which are evoked by the presentation of stimuli. Some notable examples of this category include among others:

   - **P300 ERP:** The P300 response is an Event-Related Potential (ERP), which is an electrical reaction signal detected from the nervous system subsequent to the presentation of a stimulus. It typically reaches its maximum positive value approximately 300 milliseconds after the stimulus onset, which can be tactile, visual, or auditory in nature. The amplitude of the P300 response peaks at the central and parietal regions of the scalp. It's noteworthy that the P300 response is spontaneously elicited post-stimulus application, necessitating no specific training from the subject or the user of the BCI. One popular BCI design leveraging P300 brain patterns is the P300 speller, an ERP-BCI designed for spelling different letters. In this paradigm, letters are presented randomly, and the subject is instructed to concentrate on the timing of the flashing letter corresponding to their intended spelling.

   - **Steady-State Evoked Potentials (SSVEP):** When faced with continuously flickering stimuli, the user's brain generates rhythmic patterns that mirror the frequency of the stimuli. Consequently, that translates to an elevation in EEG brain activity precisely at the stimulation frequency. BCIs utilizing this brain pattern are known for their speed and reliability and have the added benefit of requiring no training from the user or subject. One such SSVEP-BCI enables the user to control a simulated airplane, such as turning left or right, by focusing on two lights flashing at different frequencies.

However, a drawback of this approach is that the flashing lights can lead to fatigue for the user.

2. **Spontaneous:** It refers to EEG signals under a specific state without any external stimulation. For example, Sensorimotor Rhythms (SMRs) belong to this category. These brainwave patterns can be influenced by either actual movement or imagined movement, reflected in EEG signals across two frequency bands: the alpha band and the beta band. Specific activities, such as physical movement or mental imagery of movement, can prompt alterations in the power of the EEG signal. These changes in power are observable across various frequency bands depending on the tasks being performed. When power decreases, it is termed Event-Related Desynchronization (ERD), while increased power is termed Event-Related Synchronization (ERS). Alpha band Event-Related Desynchronization typically occurs a short time before the onset of real or imagined movement and peaks shortly after the initiation of the task. It returns to its baseline state after a few seconds. Conversely, beta band Event-Related Desynchronization is relatively brief at the onset of real or imagined movement, followed by Event-Related Synchronization, which reaches its maximum after the completion of the intended task.

2.3.3 EEG Signal Pre-Processing

As in any digital system, the raw signals collected during the brain activity signal acquisition step of a BCI are noisy (have low Signal-to-Noise (SNR) ratio) and contain unnecessary information for the BCI task. EEG signals measure electrical brain activity, detected by electrodes positioned on the scalp. Nevertheless, there are different layers with distinct electrical conductivity, such as the scalp, and skull bone, situated between the cortical neurons generating electrical activity and the electrodes that detect it ([38], [39]). As a result, electrical activity passing through these layers undergoes distortion before reaching the EEG sensors. This distortion introduces noise into the recorded EEG signal, complicating the recovery of the original brain activity.

In addition, EEG signals are often contaminated with artifacts. There are two main sources of these artifacts:

- **Physiological Sources:** The subject's heartbeat, an unintended motion, even a eye blink can cause an artifact in the brainwave - intrinsic artifacts.

- **Electrical Sources:** Like any electrical signal, brain activity signals can also be distorted from sources like power lines or misplaced electrodes - extrinsic artifacts.

Fortunately, motion-related physiological artifacts can be effectively removed by incorporating recordings from other sources, such as Electromyography (EMG), along with the brain activity signals. When brain activity signals exhibit strong correlation with these additional data, they are retained for the training of the BCI model.

Conversely, certain electrical artifacts can be mitigated using digital filters. For example, a Notch filter can suppress power-line interference at 50Hz or 60Hz through bandstop filtering, while a Band-Pass filter can isolate frequency content within a specific range.

Improperly positioned electrodes can lead to faulty connections, rendering the signal from these electrodes unsuitable for further analysis. Misplaced electrodes should be either removed entirely or replaced by performing interpolation from neighboring electrodes.

Additionally, to eliminate individual sources of noise, such as ocular artifacts, techniques like Independent Component Analysis (ICA) can be employed ([40], [41], [42]). First, the EEG signals are decomposed into several source components. Then, insignificant components are discarded and the retained components are used in the reconstruction of the clean EEG signals. In essence, ICA reveals hidden factors and similarities underlying the observed signal data.

Spatial filtering of brain activity signals provides the necessary quality enhancement, which is necessary for the feature extraction and classification steps of a BCI. A commonly used technique for spatial filtering of EEG signals is the Common Average Reference (CAR). Under Common Average Reference (CAR) [43], the voltage potential is averaged across all sensors (electrodes) and is subtracted from the recording electrode. Mathematically, this can be described as:

$$V_{CAR(i)} = V_i - \frac{1}{N}\sum_{j=0}^{N} V_j \tag{2.1}$$

This techniques can effectively reduce noise that is present in all EEG electrodes. When more localised filters are needed, other spatial filtering techniques techniques, like Common Spatial Patterns (CSP) [44], have been proposed.

2.4 Motor-Imagery Decoding

MI-BCIs are the primal focus of this Thesis. For someone to truly appreciate the immersive power and endless possibilities that the field of MI-BCIs unveil, they need to visualize the following scenario:

> Imagine being completely paralyzed and having lost the ability to move your hands or feet.

Motor-Imagery stands as one of the most popular BCI paradigms. MI-BCIs are based on a cognitive process involving the mental simulation of movement without actual physical execution [45]. It encompasses the vivid mental rehearsal of motor actions, engaging similar neural circuits as those involved in actual movement ([46], [47]). During Motor Imagery, individuals mentally visualize themselves performing specific movements, such as grasping an object or walking, without physically executing these actions. This mental rehearsal is believed to activate motor-related brain regions, including the primary motor cortex, supplementary motor area, and premotor cortex, as well as the same efferent pathways responsible for executing physical movements ([48], [49]). MI-BCIs are widely used for movement rehabilitation purposes (e.g., [50], [51] and [52]) - to restore and repair movement in individuals with paralysis - as well as for research purposes in cognitive psychology [53].

The decoding of EEG MI signals constitutes a challenging task due to:

1. the low Signal-to-Noise ratio (SNR) of the EEG recordings
2. the well-documented variability within-subject and across-subjects data [54]

These two major factors limit one's ability to fully understand the dynamics of the brain's activity and extract appropriate features for motor-imagery classification tasks [55].

2.4.1 Classical Decoding Techniques

Over the years, several works have addressed the problem of EEG-based motor-imagery classification using classical feature extraction techniques [56].

Among various decoding methods, Common Spatial Patterns (CSP) [44] has emerged as a prominent technique due to its effectiveness in extracting discriminative features from EEG signals.

CSP [31] has been widely adopted in MI-BCI systems owing to its ability to enhance the SNR ratio by finding spatial filters that maximize the variance difference between two classes of EEG signals. Mathematically, a spatial filter w is computed with:

$$w = argmax_w \frac{w\Sigma_1 w^T}{w\Sigma_2 w^T} \tag{2.2}$$

where Σ_1 and Σ_2 denote the averaged covariance matrix of class 1 and class 2 respectively.

Several variations of this technique have been introduced among those: [57] introduces the regularized CSP (R-CSP), a method that addresses the issue of overfitting by integrating regularization techniques into the CSP framework. Probabilistic CSP (P-CSP) [58] encompasses conventional CSP algorithms within a probabilistic modeling context. While [59] combined CSP with wavelet transform to extract spatio-spectral features.

Moreover, the integration of CSP [44] with filter bank techniques has led to the development of Filter Bank Common Spatial Pattern (FBCSP) approaches. FBCSP [60] builds upon the fundamentals of CSP, but unlike traditional CSP, FBCSP decomposes EEG signals into multiple frequency subbands using a filter bank, allowing for more precise feature extraction across different frequency bands. This multi-resolution approach enhances the discriminative power of the extracted features, enabling more robust MI-BCI decoding. While [61] introduces Common Time-Frequency-Spatial Patterns (CTFSP), which derives features from EEG data that has undergone multi-band filtering across various time segments: a sliding time window method is used to divide the EEG timeseries into multiple subseries, and sparse CSP features are extracted from different frequency bands within each of these time windows.

Common Spatial Pattern (CSP) [44] as well as its variants (especially FBCSP [60]) have gained high popularity since they are simple to understand and implement, and they are fast to run, which makes them attractive to BCI researchers. However, as described, distinct frequency bands need to be pre-defined before the actual processing of EEG signal which can be limiting, as it may not be possible to know in advance which band of frequencies will be most informative for the task at hand. Therefore, more data-driven decoding methods are needed.

2.4.2 Deep Learning Decoding Methods

With the accelerating progress in the field of Deep Learning, deep convolutional neural networks have emerged as potential alternatives, in which the exact design of the spatio-temporal filtering pipeline is accomplished in a data-driven and principled manner so as to ensure the reliability of the subsequent brain activity decoding:

- **DeepConvNet / ShallowConvNet**: DeepConvNet [62] is one of the very first neural architectures utilized in MI-related BCIs. It is a Convolutional Neural Network (CNN) and consists of four convolution-max pooling blocks. The novelty of this architecture is a specially-designed convolution in its first convolution-max pooling block, which consists of a convolution across time followed by a convolution across EEG channels. Inspired by CSP spatial filters [44], the filters in this block have weights for all EEG channels. A shallower deep BCI architecture is ShallowConvNet [62]. Inspired by the FBCSP [60], it consists of two convolutions (one temporal and one spatial), a squaring non-linearity, a max pooling operation and a logarithmic non-linearity.

- **EEGNet:** A notable lightweight BCI architecture is EEGNet [63]. It consists of fully trainable convolutional temporal filters shared across all EEG channels followed by spatial filters formed by linear combinations among the channels. The compound of temporal and spatial filtering was inspired by the FBCSP technique [60].

- **EEG-Inception**: Another state-of-the-art BCI architecture with strong performance results across different benchmarks is EEG-Inception [64] which follows the exact same principles as EEGNet [63] but includes the notion of EEG-multiscalability. The EEG-Inception architecture is similar to EEGNet [63] but additionally includes several Inception modules, originally introduced in [65]. These modules consists of trainable convolutional temporal filters of different scales, capturing several temporal modulations of the EEG signals. The ability of EEG-Inception to incorporate brain dynamics spanning distinct temporal scales was considered of importance in the MI case, and hence included in the present Thesis as a state-of-the art technique.

2.5 Motor-Imagery Datasets

In the experimental sections of this Thesis, multichannel EEG signals from three publicly available MI datasets, namely PhysioNet [66], BCI IVa Dataset of the 3rd BCI competition [67] and OpenBMI - MI [68], were used so as to train and test various deep learning models of MI-recognition, both at population and personalised level. Single-trial EEG segments were extracted (based on the timing of delivered visual cues) and were treated as composite patterns for a classification problem of MI-recognition. The three datasets differed in many ways (e.g. sampling frequency and sensor configuration) and treated differently during the pre-processing, as briefly described below:

1. **PhysioNet**: The original dataset consists of brain recordings from 109 healthy participants, registered via 64 EEG sensors (BCI2000 system [3]) with a sampling frequency of 160 Hz, while performing a series of pseudo-randomized cue-triggered MI tasks. The subjects performed either real or imagined movements based on a visual stimulus:

   (a) If a target appeared on either the left or right side of the screen, the subject clenched (or imagined clenching) and released the corresponding fist until the target disappeared.

   (b) If a target appeared on either the top or bottom of the screen, the subject clenched (or imagined clenching) and released both hands (top) or both feet (bottom) until the target disappears.

   In our experiments, 6 participants (subjects 88, 89, 92, 100, 104 and 106) were excluded from the dataset due to differences in either the sampling frequency or duration of the performed tasks. Before extracting trials from these continuous mode recordings, a series of "data-cleaning" steps was performed so as to reduce the effects of contamination by physiological artifacts. This step was dictated by the fact that only a small number of trials was available for each participant, and we had to ensure sufficient information for training (and validating) personalised models. The pre-processing included the following steps:

   - A Bandpass Filter with 3rd Order Butterworth (applied in zero-phase mode) at 0.5 - 79.5 Hz
   - Notch Filter at 60Hz for powerline noise removal

[3]http://www.bci2000.org

- Common Average Re-referencing: subtracting the mean of the channels at each time-point from each channel
- Application of an in-house implementation of w-ICA (Independent Component Analysis) [69] as described in [70].

Finally, we extracted trials corresponding to MI hand or feet movements in the form of segments starting with the visual cue and lasting for 4.1 seconds.

BCI IVa Dataset: The original dataset includes brain recordings from 5 healthy subjects, registered via 118 EEG sensors, while performing a series of randomized cue-triggered MI tasks. Trials corresponding to MI of right hand or right foot were extracted from the continuous signal in the form of segments starting with the visual cue and lasting for 3.5 sec. There were 280 trials per subject, sampled at 100 Hz, that served as a benchmark dataset for our experiments without any pre-processing.

OpenBMI - MI [68]: The original OpenBMI dataset consists of 3 BCI paradigms: ERP-based speller, MI and SSVEP. The MI trials include brain recordings from 54 healthy participants, registered via 62 EEG sensors with a sampling frequency of 1000 Hz. In the MI part of the dataset, the participants performed a series of cue-triggered MI tasks, either with or without receiving feedback (cursor moved according to the prediction of a trained classifier) : each trial commenced with a black fixation cross displayed at the center of the monitor for the first 3 seconds, serving to prepare subjects for MI task. Following this, the subject engaged in the imagery task of grasping with the corresponding hand for a duration of 4 seconds upon the presentation of a right or left arrow as a visual cue. Subsequently, after each task, the screen remained blank. For the purpose of this Thesis, we kept only the MI-trials without feedback. In particular, we extracted trials corresponding to MI hands in the form of segments starting with the visual cue and lasting for 4 seconds. Furthermore, we applied a notch filter at 60Hz - and its harmonics (120, 180, 240, 300, 360, 420, 480) - to remove powerline noise. We also applied a notch filter at 460Hz due to a spurious artifact (consistent across all trials).

3

Capturing Real-World Brainwave Data

Contents

In this chapter, we describe the data collection experiment, titled: *"**MyBrainCommands**[4]: Mind Reading for control"*, that took place at the Department of Computing at Imperial College London. All the participants gave written consent. The experiment received approval by the Head of Department and the Research Governance and Integrity Team (SETREC number 22IC7505), 2022-2023. It should be noted that this data collection constitutes *joint work* with other collaborators at Imperial College London.

[4]Website: https://mybraincommands.doc.ic.ac.uk

3.1 Introduction

Designing EEG decoding models using deep learning methods relies heavily on the availability of datasets, which pose a significant challenge due to their scarcity and limited accessibility. EEG datasets, crucial for training and validating these models, are predominantly sourced from specialized laboratories or hospital settings. As a consequence, obtained from these controlled environments, these datasets often fail to represent real-world scenarios adequately. While they provide valuable insights into brain activity under controlled conditions, they lack the diversity and complexity necessary to address the myriad of real-world applications for EEG decoding. As it is evident, we are in need of more EEG datasets that

closely mirror **real-world** conditions beyond the confines of the laboratory

The MyBrainCommands experiment at Imperial College London was crafted with the aim of addressing this need. The general study aimed at recording and modeling brain activity in various real-life situations based on wireless recording devices. The experimental protocol had been approved by the local ethics committee at Imperial College London and the subjects signed a consent form.

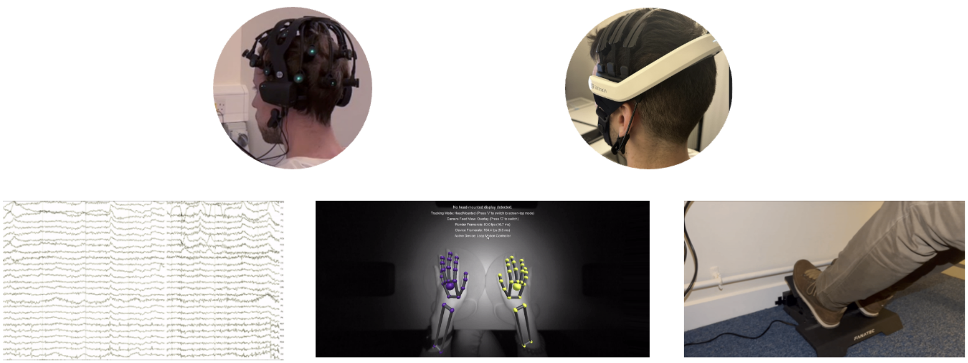

Figure 3.1: MyBrainCommands Setup - (Top) The two EEG headsets that were used in the experiment. (Bottom) The three modalities captured in the experiment: EEG, hand movement, feet movement.

In this experiment, we recorded brainwaves of participants while they were engaging in simple movement tasks, namely clenching their fists (real and imagined movements) pressing pedals with their feet (real and imagined movements). Three modalities were captured: EEG brainwave data, hand movement data (using hand tracking camera[5]) and feet movement data (using pedals[6]). Two wireless dry recording devices were used to collect the EEG data, namely Quick 32r Cognionics device[7] and Hero Bitbrain device[8].

3.2 Dataset Statistics

At the time of composing this Thesis, 504 unique recording sessions have been collected from 332 unique participants. The age range of the participants spanned from 19 to 80 years, with a mean age of 28 and a standard deviation of 9. The comprehensive statistics of the participants can be found in following figure:

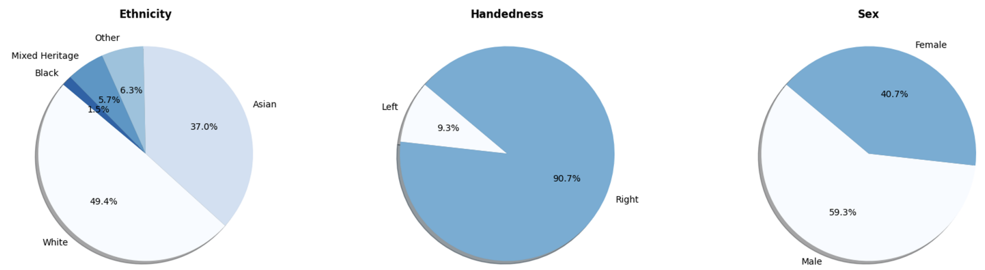

Figure 3.2: Statistics of the participants that took part in the MyBrainCommands experiment.

The experiment has been divided into three paradigms, namely Standard Setting, Gamified Setting and Self-Paced Setting, which will be described in the next sections.

Paradigm	Setting	EEG Recording Device Used	Number of Recordings
Standard Setting	Standard Lab	Quick 32r Device	183
	Standard DoubleClick	Quick 32r Device	40
	Standard DoubleClick	Hero Bitbrain Device	51
Gamified Setting	Robot Factory	Quick 32r Device	112
	Robot Factory DoubleClick	Quick 32r Device	72
	Robot Factory DoubleClick	Hero Bitbrain Device	46
Self-Paced Setting	Racing Game	Quick 32r Device	295

Table 3.1: Number of recodings for each paradigm of the MyBrainCommands dataset.

[5] https://www.ultraleap.com/leap-motion-controller-whats-included/
[6] https://fanatec.com/eu-en/csl-elite-pedals
[7] https://www.cgxsystems.com/quick-32r
[8] https://www.bitbrain.com/neurotechnology-products/dry-eeg/hero

3.3 Standard Setting

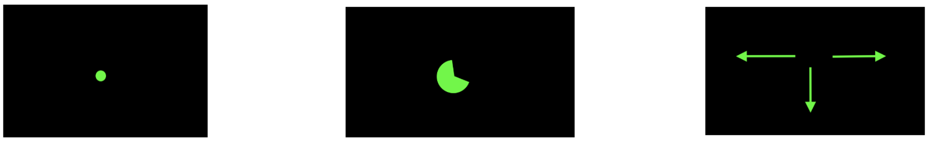

Figure 3.3: MyBrainCommands Standard Setting - (Left) Visual cue indicating the beginning of a trial. (Middle) Countdown timer to distinguish the stimulus-related ERP response of the brain from the motor-related brain activity. (Right) Target arrows indicating the movements.

The Standard Setting of the MyBrainCommands experiment is a lab-structured environment where the user should perform specific time-locked movements (real or imagined) based on some visual cues in the form of target arrows on the screen. This setting was split into two variations, namely Standard Lab Setting and Standard DoubleClick Setting.

3.3.1 Standard Lab Setting

Each trial commences with a blank screen and a visual cue in the middle of the screen indicating the beginning of the trial. Subsequently, a visual cue indicates the movement (real or imagined) that the subject should execute. The subject then performs the movement (real or imagined) after a countdown timer (preparation time) has elapsed to distinguish the stimulus-related ERP response of the brain from the motor-related brain activity. The subject should perform one of the following movements:

1. An arrow is displayed pointing to either the left or right side of the screen. The subject clenches (in the real setting) the corresponding fist (after the countdown) and releases when the cue disappears. Then, the subject relaxes.

2. An arrow is displayed pointing to either the left or right side of the screen. The subject mentally visualizes clenching (in the imaginary setting) the corresponding fist (after the countdown) and releases when the cue disappears. Then, the subject relaxes.

3. An arrow is displayed pointing to the bottom of the screen. The subject presses both feet (in the real setting) and releases when the cue disappears. Then, the subject relaxes.

4. An arrow is displayed pointing to the bottom of the screen. The subject mentally imagines pressing both feet (in the imaginary setting) and releases when the cue disappears. Then, the subject relaxes.

3.3.2 Standard DoubleClick Setting

Each trial commences with a blank screen and a visual cue in the middle of the screen indicating the beginning of the trial. Subsequently, a visual cue indicates the movement (real or imagined) that the subject should execute. The subject then performs the movement (real or imagined) after a countdown timer (preparation time) has elapsed to distinguish the stimulus-related ERP response of the brain from the motor-related brain activity. The subject should perform one of the following movements:

1. An arrow is displayed pointing to either the left or right side of the screen. The subject clenches and releases the corresponding fist twice in quick succession (after the countdown). Then, the subject relaxes.

2. An arrow is displayed pointing to either the left or right side of the screen. The subject mentally visualizes clenching and releasing the corresponding fist twice in quick succession (after the countdown). Then, the subject relaxes.

3. Two arrows are displayed pointing to both sides (left and right). The subject clenches and releases both fists twice in quick succession. Then, the subject relaxes.

4. Two arrows are displayed pointing to both sides (left and right). The subject mentally imagines clenching both fists twice in quick succession. Then, the subject relaxes.

5. An rectangle is displayed. The subject should perform no movement (real or imagined). Then, the subject relaxes.

3.4 Gamified Setting

The Gamified Setting of the MyBrainCommands experiment is a structured gamified environment where the user should perform specific time-locked movements (real or imagined) based on some visual cues in the form of robots in a game. This setting was split into two variations, namely Robot Factory Setting and Robot Factory DoubleClick Setting.

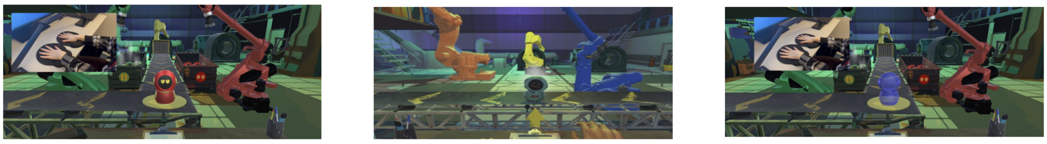

Figure 3.4: MyBrainCommands Gamified Setting - (Left) Visual prompt of red robot. (Middle) Visual prompt of grey robot. (Right) Visual prompt of blue robot.

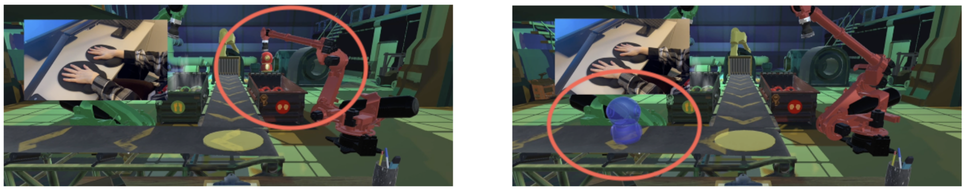

Figure 3.5: MyBrainCommands Gamified Setting - (Left) Game's response to a movement-related robot. (Right) Game's response to a non-movement-related robot.

3.4.1 Robot Factory Setting

Each trial commences with a game factory setting displayed on the screen. Subsequently, a robot arrives (visual cue) with its color indicating the movement (real or imagined) that the subject should execute. The subject then performs the movement (real or imagined) once the robot stops in front of the user (preparation time) to distinguish the stimulus-related ERP response of the brain from the motor-related brain activity. The subject should perform one of the following movements:

1. A red robot arrives. The subject clenches (in the real setting) and releases the right fist when the robot is collected by a crane. Then, the subject relaxes.
2. A red robot arrives. The subject mentally imagines clenching the right fist (in the imaginary setting) and releases when the robot is collected by a crane. Then, the subject relaxes.
3. A grey robot arrives. The subject clenches (in the real setting) and releases the left fist when the robot is collected by a crane. Then, the subject relaxes.
4. A grey robot arrives. The subject mentally imagines clenching the left fist (in the imaginary setting) and releases when the robot is collected by a crane. Then, the subject relaxes.
5. A blue robot arrives. The subject should perform no movement (real or imagined). Then, the subject relaxes.

3.4.2 Robot Factory DoubleClick Setting

Each trial commences with a game factory setting displayed on the screen. Subsequently, a robot arrives (visual cue) with its color indicating the movement (real or imagined) that the subject should execute. The subject then performs the movement (real or imagined) once the robot stops in front of the user (preparation time) to distinguish the stimulus-related ERP response of the brain from the motor-related brain activity. The subject should perform one of the following movements:

1. A red robot arrives. The subject clenches and releases (in the real setting) both fists twice in quick succession. Then, the subject relaxes.

2. A red robot arrives. The subject visualizes clenching (in the imaginary setting) and releasing both fists twice in quick succession. Then, the subject relaxes.

3. A blue robot arrives. The subject should perform no movement (real or imagined). Then, the subject relaxes.

3.5 Self-Paced Setting

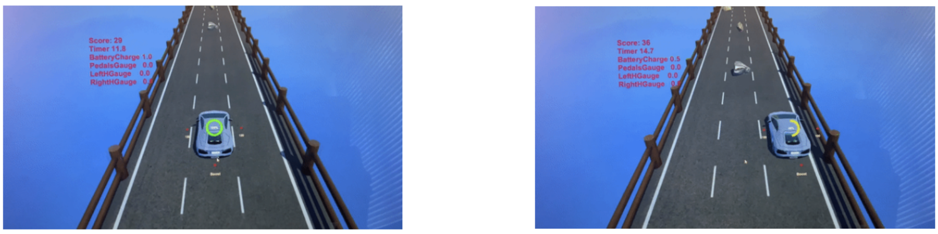

Figure 3.6: MyBrainCommands Self-Paced Setting. - (Left) Car accelerates after the successful execution of legs' pressing. (Right) Car turns right after the successful execution of a right-hand movement.

In the Self-Paced Setting of the MyBrainCommands experiment, users engage in a gamified environment where they control a car in a racing game by planning and executing real movements. There are no on-screen instructions. The aim is to travel as far as possible while avoiding rocks on the road. The subject should perform one of the following movements:

1. The subject clenches and releases the right fist to change to the adjacent right lane of the road.

2. The subject clenches and releases the left fist to change to the adjacent left lane of the road.

3. The subject presses and releases both feet to speed up the car for a short amount of time.

The game is controlled using the hand motion and pedal devices. Additionally, there is a countdown following each car movement, preventing the player from making any further movements during this time interval. This measure aims to prevent the blending of multiple motor-related brain activities and movements (co-activations).

3.6 Conclusion

In this chapter, we described the "MyBrainCommands: Mind Reading for control" experiment - a very recent large-scale EEG data-collection at the Department of Computing at Imperial College London. This dataset includes standard lab setting paradigms that provide a structured framework for EEG data collection, allowing us to examine brain responses under specific stimuli or tasks. It also extends into self-paced gamified environments, incorporating an innovative dimension to EEG data collection. By integrating elements of gaming into the experimental protocol, we aimed to engage participants in more naturalistic and immersive experiences, not only enhancing participant motivation and enjoyment but also eliciting more dynamic brain signals.

In the following chapters (i.e, Chapter 5), the subset Standard Lab Setting - MI was utilized in the experimental section. The particular subset of the in-house dataset concerns brain activity recorded, via Quick 32r headset, from adult volunteers while performing a series of randomized cue-triggered MI tasks. As described, participants performed imaginary closing of left fist 2.5 seconds after a visual prompt (countdown time) had been flashed on the screen. Trials corresponding to MI of left fist were extracted from the continuous signal in the form of segments starting at the expected movement onset and lasting for 2.5 sec. The post-movement rest interval had been randomized at 3 ± 0.2 sec and its initiation was prompted visually (by the disappearance of the cue). Brain activity segments, 2.5 seconds in duration, were extracted from this interval and used to cast a binary classification task of the form of "Left Hand" vs "Rest" that would enable training a model for facilitating a "brain-switch" [71]. There were 27 trials per class per subject and since we wanted to approximate a real-life application in our experiments, our pre-processing included the following steps: 1) highpass filtering at 0.5 Hz, 2) notch filter at 60Hz for powerline noise removal, 3) common average re-referencing and signal resampling at 160Hz and 4) frontal electrodes were considered susceptible to ocular artifacts and left out of the modeling, hence the utilized sensors were: Fz, FC6, F4, C4, Oz, CP6, Cz, PO8, CP5, O2, O1, P3, P4, P7, P8, Pz, PO7, C3, F3, FC5.

4

A Causal Perspective on Brainwave Modeling for BCIs

Contents

In this chapter, we introduce a framework aiming to breakdown and analyze important challenges of brainwave modeling for BCIs based on Causal Reasoning, which was published in **K. Barmpas**, Y. Panagakis, G. Zoumpourlis, D. Adamos, N. Laskaris, and S. Zafeiriou, *"A Causal Perspective on Brainwave Modeling for Brain-Computer Interfaces"*, in Journal of Neural Engineering (2024).

4.1 Introduction

BCIs have the potential to revolutionize the way we interact with the world around us. In the future, BCIs could be used to help people with disabilities regain lost functions, enhance our cognitive abilities, and create new forms of entertainment and communication ([72], [73], [74], [75]). The possibilities for BCIs are endless and their future heavily relies on developing techniques and models that can accurately analyze the captured brainwave signals.

As described in Chapter 2 in detail, for several years, neuroengineers have been trying to analyze EEG signals using classical signal processing methods ([56], [76], [77], [78], [79]) and extracting manually crafted features, making the "expert's knowledge" the center and a vital part of brainwave analysis. In recent years, however, the field of brainwave modeling has been enjoying tremendous progress with the advent of Machine Learning and Deep Learning ([80], [81], [82]), which has alleviated the need for manual feature extraction and has provably revolutionized several other fields like Computer Vision (e.g. [83] and [84]), Audio / Speech Recognition (e.g. [85]) and Natural Language Processing (e.g. [86] and [87]). The application of deep neural networks in various paradigms in the area of BCIs - e.g. epilepsy detection and prediction ([88], [89]), sleep stage detection [90], anomaly detection [91], muscle activity [92], acute stress detection [93], and motor-imagery classification [94] - has achieved state-of-the-art performances compared to previously employed classical signal processing techniques.

However, several obstacles in claiming a wide range success of Machine Learning in the field of BCI still remain: the limited amount of available annotated data required to train these networks as well as the quality (data noise) and fundamental differences (different recording conditions and various demographics) between the available brainwave datasets. Let us illustrate with a hypothetical case scenario how these challenges may arise in practice and pose real threats to various BCI paradigms. Suppose that a university laboratory team has collected some MI data from several students (similar to MyBrainCommands as described in Chapter 3). The laboratory team used these collected data to train an MI-BCI to decode brainwaves. The team was able to demonstrate high performance in this task that outperformed classical signal processing methods as confirmed via the ground-truth MI labels. Because of their success, they hope to license their model to rehabilitation centers to assist in the restoration of motor function in individuals with paralysis.

And the question arises, is this project bound to succeed or fail? There are various issues and challenges that can jeopardize its success which might include: the limited amount of annotated data, data mismatch between the training and deployment sets, different population groups (the current dataset contains only young students which might not be the case in the rehabilitation center) and different EEG headsets used for the data collection and for the rehabilitation tasks.

In this chapter, the importance of causal relationships between the variables of interest in various tasks of brainwave modeling in the design of machine learning models in the field of BCI is discussed. In summary, this chapter's contributions are as follows:

- Causal Reasoning is employed to accurately describe the task of brainwave modeling using independent causal factors
- A causal framework is introduced that allows all BCI paradigms to be categorized based on two core properties: the presence of task stimulus and the voluntary engagement of the subject
- Solutions to the identified challenges through this causal factorization are described, which can be either general machine learning practices or domain-specific techniques
- Evaluation schemes are discussed to efficiently compare brainwave modeling methods in the present and the future

The remainder of the chapter is organized as follows: Section 4.2 describes related work in the literature, details the motivation behind our causal framework and lies down the theoretical background of Causal Reasoning that is employed in the next sections of this chapter. Section 4.3 outlines our causal framework. More specifically, it categorizes brainwave models to encoders and decoders and breaks them down to their independent causal factors. Section 4.4 describes some of the challenges that are met in brainwave modeling based on this causal analysis and outlines techniques that can be employed to solve these identified obstacles. Section 4.5 portraits appropriate evaluation schemes. The last section 4.6 summarizes and concludes the work presented in this chapter.

Inspired by [95] in the field of Medical Imaging, the main objectives in this chapter are:

1. to **understand** the problem of brainwave decoding and
2. to raise awareness that "causality **matters** in brainwave analysis"

It is hoped that this causal framework will serve as a guideline for future research in the field of brainwave modeling, leading to massive advances as well as successful commercialization of BCIs.

4.2 Preliminaries in Causal Reasoning

4.2.1 From Physical to Statistical Models

Our real world constantly evolves over time and we are usually forced to adapt, either actively (by choice) or passively, to these changes. Modeling reality has been an important goal for humankind and the aspiration behind the invention of many modern scientific tools. Physical phenomena and physical systems were among the first mechanisms that scientists managed to successfully model using mathematics.

A commonly-used mathematical tool is differential equations, for example motion equations, providing a comprehensive description of physical/mechanical systems while capturing their underlying casual structures. Reliable predictions under different changes to their variables is a feature that all physical models share. Despite their provable robustness to changes or shifts that can take the model outside of its original setting, they usually require a human expert to understand the system that needs to be modelled (its structure, its dynamic processes, the causal relationships between its different parts, etc.) and come up with these analytical equations.

For certain systems or problems, though, it is often impossible to reach analytical equations that accurately describe them, no matter the available expert knowledge at hand. In these cases, statistical models (widely used in the field of Machine Learning) are proved to be useful. This type of models heavily relies on observational data and can uncover associations between the input and output data. In these models, it is possible to infer some accurate predictions based only on observational input data as long as there are no changes in the system or the problem that is attempted to be modeled.

Models learnt using machine learning algorithms belong to the category of statistical models, since these algorithms try to make accurate predictions from a given amount of observations. Nowadays, the great success of Machine Learning (or Deep Learning) in various fields is the result of four main factors, as described in [96]:

1. large amount of data
2. massive computational power
3. high-capacity systems
4. independent and identically distributed (i.i.d.) tasks

Today's hardware and research practices guarantee almost always factors **2.** and **3.** On the other hand, factor **1.** tends to be valid for the majority of the problems tackled by machine learning models, although there are still areas where a lengthy data acquisition process and/or lack of human annotations prohibit the availability of large amounts of (labelled) data for particular tasks.

Lastly, the i.i.d. assumption (factor **4.**) seems to be the greatest vulnerability for the majority of machine learning models. When this assumption is violated, by introducing changes or distribution shifts that take the models outside of their current settings, it is empirically proven that these algorithms are bound to fail (tend to not generalize well - in machine learning terms). These distribution shifts vary from simple perturbations to the input data (adversarial samples [97]) to different data distributions being used during development and deployment.

4.2.2 Causal Models

Unlike physical models, statistical models are not always robust to distribution shifts when at least one of the previously mentioned factors is violated. Generalization across problems and distribution shifts is an open challenge in the machine learning community and a gap that Causal Reasoning promises to bridge.

Causal Reasoning is the analysis of a task/problem in terms of cause-effect relationships between the different variables of interest [95]: if a variable A is a direct cause of variable B, it is expressed as $\mathsf{A} \to \mathsf{B}$ (causal diagram: A causes B or B is the effect of A).

Causal modeling lies somewhere in-between these two types of models (physical and statistical). On the one hand, it takes into account causal relationships between the different variables of interests, like physical models, and on the other hand it is data-driven (learnt from observations), like statistical models. Figure 4.1 graphically demonstrates the difference between a statistical and a causal model.

Given an input X and a target prediction Y, a statistical model tries to estimate $P(Y|X)$. A causal model, although data-driven, extends the statistical setting by considering additionally the cause-effect relationships between the various variables of interests. Using causal terminology, a task can be either [95]:

- Causal: when $X \to Y$, prediction of effect from cause or
- Anti-causal: when $Y \to X$, prediction of cause from effect

According to the Principle of Independent Causal Mechanisms (ICM) ([96], [98], [99]), the causal breakdown of a system consists of independent variables that do not influence or inform each other[9].

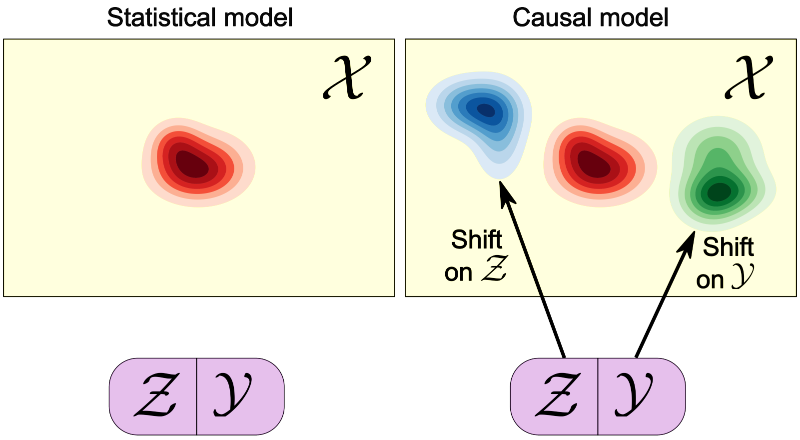

Figure 4.1: Statistical (left) and causal models (right) on a given set of two variables Z and Y. In the case of statistical model a single probability distribution is specified while in the case of causal model a set of distributions, one for each possible intervention, is specified - Figure inspired by [96].

In order to demonstrate the power of causal models, let us consider the following scenarios where we have two observable variables X and Y. According to the Common Cause Principle (Reichenbach's Principle)[10], the following scenarios [96] exist:

1. $\mathsf{X} \to \mathsf{Y}$ where X causes Y or Y is the effect of X
2. $\mathsf{X} \leftarrow \mathsf{Y}$ where Y causes X or X is the effect of Y
3. $\mathsf{X} \leftarrow \mathsf{Z} \to \mathsf{Y}$ where both variables X and Y are caused by another variable Z

A statistical model would not be able to differentiate between these three cases, simply due to the fact that the joint observable distribution $P(\mathsf{X}, \mathsf{Y})$ is the same in all three cases. By taking into account, the causal relationships between the variables of interest, a causal model has automatically more information - which in turn if appropriately used by machine learning algorithms - can lead to models which are more robust to certain types of distribution shifts.

[9] A causal breakdown of a system can be represented as a Directed Acyclic Graph (DAG) where the nodes are the variables of interests and the edges represent direct causal relationships [11].

[10] Common cause principle ([100], [96]): If two observable variables X and Y are statistically dependent, then either there is causal relationship between them or there exists a third variable Z that causally influences both and makes them independent when conditioned on Z.

4.3 Causality in Brainwave Modeling

4.3.1 Brainwave Models

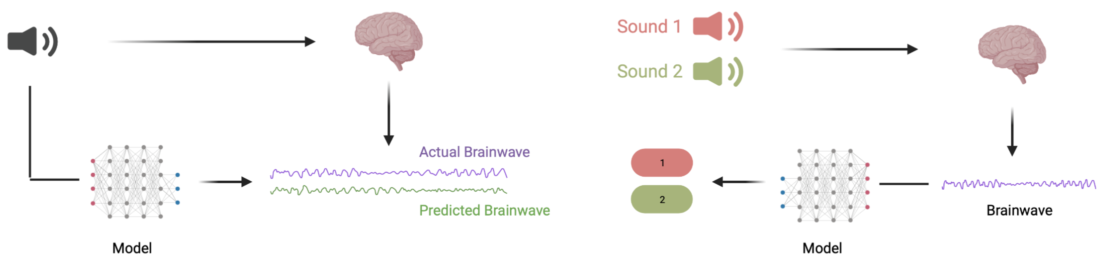

Figure 4.2: (Left) An example of an encoding brainwave model that predicts the brainwave activity based on a stimulus - (Right) An example of a decoding brainwave model that predicts the label based on the brainwave activity.

Understanding the brain is arguably one of the most interesting problems of our time. EEG technology has made it possible for researchers to measure the brain activity of a subject in a non-invasive manner, while various tasks are performed. Since the underlying organizational structure of the brain still remains a puzzle to be solved, there is no analytical model (physical modeling) that can accurately describe its functionality for the entirety of human behavior[11]. Therefore, the success of BCIs mainly relies on the recent advances in Machine Learning (statistical models).

BCI models can be distinguished between encoding and decoding models (examples are shown in Figure 4.2), known in machine learning terms as generative and discriminative models:

- Encoding models : Given a known performed task/ stimulus Y, the goal of these models is to estimate the brain activity X (brainwave) corresponding to that particular task, i.e. $P(X|Y)$
- Decoding models: Given an observed brain activity X (brainwave), the goal of these models is to predict the performed task Y corresponding to that particular brainwave, i.e. $P(Y|X)$

By introducing causal (cause-effect) relationships between the observed brain activities and their corresponding performed tasks, these two types of brainwave models can be identified with contrasting causality terms:

[11] It is worth mentioning here an increasing amount of effort to design tools that simulate the brain using methods from statistical physics like The Virtual Brain (TVB) [101]

- If $X \rightarrow Y$: Encoding models are anti-causal and decoding models are causal
- If $Y \rightarrow X$: Encoding models are causal and decoding models are anti-causal

Although encoding BCI models (BCI systems that modulate the current neural activity so as to bring it within a desirable range associated with some targeted label/ state) can be met in real-life application like transcranial direct current stimulation for limb rehabilitation [102] or Parkinson's disease [103], the vast majority of BCI systems are decoding models, therefore we will focus only on BCI brainwave decoders.

4.3.2 Causal Framework

To make use of causality models, as described in section 4.2, it is crucial to break down the task of brainwave analysis in a number of independent variables, according to the Principle of Independent Causal Mechanism, and determine the underlying causal relationships between them. To effectively perform this analysis, the various tasks of BCI paradigms need to be first categorized. Unlike previous research works [104], we argue that this categorization is bound to two core properties - the presence of task stimulus and the voluntary engagement of the subject.

Based on the presence or absence of stimulus, a BCI paradigm can be identified as:

- Exogenous: when the brain activity is modulated by the presence of stimuli which can take any form of sensory input.
- Endogenous: when the brain activity is inherently modulated, without the intervention of any external stimulus.

On the other hand, based on the subject's voluntary engagement, a BCI paradigm can be identified as:

- Voluntarily Engaged: When the subject willingly generates a particular brain activation or coactivation pattern like in the case of Motor Imagery [105].
- Involuntarily Engaged: When the subject has no control of the generated brain activation pattern. In other words, the ongoing brain activity is modulated without any conscious effort from the subject [106].

Therefore, all BCI paradigms can be categorized using a combination of the above mentioned categories. For example, P300-Speller [107] (a type of Brain-Computer Interface that allows users to spell words and sentences by focusing their attention on specific characters) is exogenous (since there is a visual stimulus) and voluntarily engaged (since the subject initiates the brain activity). On the other hand though, P300 Lie Detector [108] (a type of Brain-Computer Interface that uses ERPs to detect deception) is again exogenous and involuntarily engaged (since the subject has no control over the generated brain patterns from the visual stimulus). In Table 4.1, examples of some BCI paradigms based on this causal framework are demonstrated.

(Ex)-(VE)	(Ex)-(IE)	(End)-(VE)	(End)-(IE)
P300 speller [107] SSVEP [109] cVEP [113] Auditory attention [117]	P300 lie detector [108] Error-related potentials [110] Object recognition ERPs [114] Music appraisal [118]	Motor imagery [94] Executed movement [111] Neurofeedback [115]	Seizure detection [88] Workload estimation [112] Sleep staging [116] Omitted target [119]

Table 4.1: Examples of BCI paradigms categorized based on our causal framework. (Ex) represents Exogenous, (End) represents Endogenous, (VE) stands for Voluntarily Engaged and (IE) stands for Involuntarily Engaged.

In this causal analysis of all BCI paradigms, the causal relationship between the observed EEG brainwave signal X and the labeled task Y is of primary interest. However, before each of these four types of BCI paradigms is analyzed separately in terms of their independent components and causal relationships, a variable shared by all categories will be introduced: the true underlying unobserved BCI-relevant brain activity Z. Inspired by [95], the EEG signal X can be considered a noisy measurement of the true unobserved neural activation patterns Z, i.e., $\mathsf{Z} \to \mathsf{X}$ ([11], [12]).

Given a BCI brainwave decoder that is exogenous, when a stimulus S is applied, there are two possible scenarios:

1. The task Y causes the stimulus S which in turn generates the corresponding brain activation pattern Z contained in the measured signal X (exogenous and voluntarily engaged). In this case, the causal diagram of the variables of interest is: $\mathsf{Y} \to \mathsf{S} \to \mathsf{Z} \to \mathsf{X}$

2. The task Y and the stimulus S together - as a pair - cause the generation of the corresponding neural activation pattern Z measured within X (exogenous and involuntarily engaged). In this case, the causal diagram of the variables of interest is: $(\mathsf{Y}, \mathsf{S}) \to \mathsf{Z} \to \mathsf{X}$

For example, in the case of P300 speller, the user chooses which letter they want to focus on and in turn that letter causes the right stimulus which results in the observed EEG brain signal. On the other hand, in the P300 lie detector case, the participant-suspect is presented with a crime-related image (stimulus) and depending on whether he/she is aware of it or not, the P300 response may involuntarily be emitted. Given these two causal diagrams, in both cases, a decoding brainwave model of this category is anti-causal.

When the system is endogenous, there is no stimulus S applied that can initiate any task. Therefore, further investigation is needed and determine the true source of the neural activation pattern Z as well as the causal sequence of all involved modules.

In the case of endogenous and involuntarily engaged BCIs, no stimulus can initiate a task's brain activity and at the same time the subject has no control over the generated brain signals. Therefore, in order to determine the brainwave source in this case, two different BCI examples of this category will be described in detail, namely epileptic seizure detection ([88], [89]) and driver's drowsiness detection [120]. In the first example, a brainwave decoding model needs to be able to distinguish between normal and epileptic brain activity. In the second example, a brainwave decoding model needs to be able to distinguish between the brain activity of alertness and drowsiness. In both cases, the true underlying brain activity changes by factors that are unknown to and uncontrollable for the subject and essentially affect their current state. As a result, in this category the environment E variable will be introduced. The changes of E variable determine the task label Y (e.g. normal or epileptic brain activity) which in turn generates the corresponding neural activation pattern Z measured by X, i.e. $\mathsf{E} \rightarrow \mathsf{Y} \rightarrow \mathsf{Z} \rightarrow \mathsf{X}$. Given this causal diagram, a decoding brainwave model of this category is anti-causal.

Finally, in the case of endogenous and voluntarily engaged BCIs, producing the causal diagram is not a trivial task to perform. Here, no stimulus can initiate the brain activity of a task, but the subject has full control over the generated brain signals. Therefore, the intention I of the subject, can be considered the source of the BCI-relevant neural activation pattern. In this case though, different parts of the brainwave sequence can be used during modeling. Let *Dec* be a decoding model that uses brainwaves to discriminate between executed left and right hand movements. If *Dec* uses the brain signal during the execution of the movement, the intention has caused the prediction label Y which causes the brain activity, i.e. $\mathsf{I} \rightarrow \mathsf{Y} \rightarrow \mathsf{Z} \rightarrow \mathsf{X}$. If *Dec* uses the motor preparation brain signal X [121], then the signal precedes the final result (the movement execution) but still the prediction label Y of the pre-movement signal is the cause of the brain activity, i.e. $\mathsf{I} \rightarrow \mathsf{Y} \rightarrow \mathsf{Z} \rightarrow \mathsf{X}$. In both cases, there is one common causal diagram and the decoding model is anti-causal.

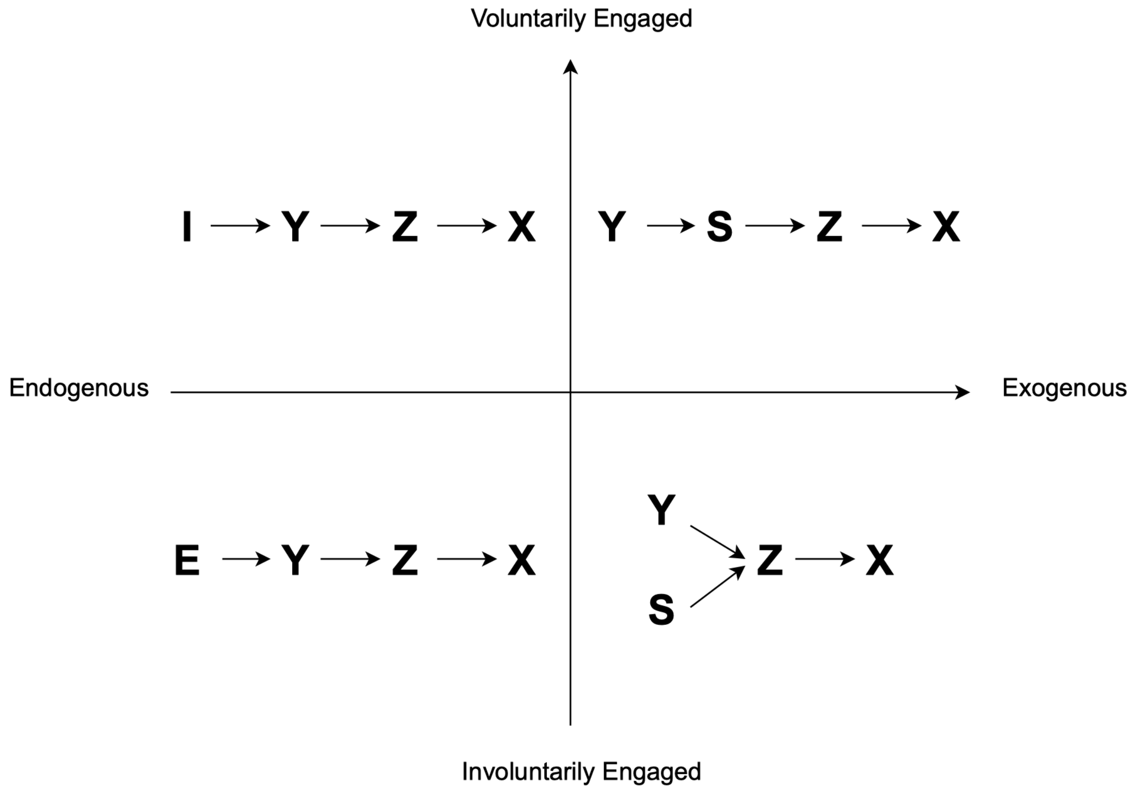

Figure 4.3: Our causal framework used to breakdown the task of brainwave decoding into causal sub-modules. For each category, there is a specific directed acyclic graph (DAG) that represent the causal relations between the identified variables of interests: (Y) represents the task true label, (S) represents the stimulus, (I) the subject's intention and (E) the environment factor.

The causal diagrams of Figure 4.3 show how different factors can influence brainwave activity. Based on these diagrams, causal factorizations can be developed, which are mathematical models that represent the relationships between these factors, to further explore the problem of brainwave modeling. For convenience, only the case of brainwave decoding models will be considered. In other words, given an input EEG signal X we want to predict the associated task Y i.e $P(\mathsf{Y}|\mathsf{X})$. [12]

- Voluntarily Engaged - Exogenous ($\mathsf{Y} \rightarrow \mathsf{S} \rightarrow \mathsf{Z} \rightarrow \mathsf{X}$):

$$\begin{aligned} P(\mathsf{Y},\mathsf{S},\mathsf{Z},\mathsf{X}) &= P(\mathsf{Y})P(\mathsf{S}|\mathsf{Y})P(\mathsf{Z}|\mathsf{S},\mathsf{Y})P(\mathsf{X}|\mathsf{Z},\mathsf{S},\mathsf{Y}) \\ &= P(\mathsf{Y})P(\mathsf{S}|\mathsf{Y})P(\mathsf{Z}|\mathsf{S})P(\mathsf{X}|\mathsf{Z}) \end{aligned} \tag{4.1}$$

[12] A similar causal diagram can also be drawn in the case of brainwave encoders where the current brain activity is modulated to a desired waveform based on targeted state. Given the current EEG signal X_{now} and the desired state Y, a stimulus S modulates the subject's brain activity Z to a desired state measured by $\mathsf{X}_{modulated} \equiv \mathsf{X}$, i.e $(\mathsf{X}_{now}, \mathsf{Y}) \rightarrow \mathsf{S} \rightarrow \mathsf{Z} \rightarrow \mathsf{X}$. Since we are now interested in $P(\mathsf{X}|\mathsf{Y})$, a brainwave encoder is a causal model.

- Involuntarily Engaged - Exogenous ($(\mathsf{Y},\mathsf{S}) \rightarrow \mathsf{Z} \rightarrow \mathsf{X}$):

$$\begin{aligned} P(\mathsf{Y},\mathsf{S},\mathsf{Z},\mathsf{X}) &= P(\mathsf{Y},\mathsf{S})P(\mathsf{Z}|\mathsf{Y},\mathsf{S})P(\mathsf{X}|\mathsf{Z},\mathsf{Y},\mathsf{S}) \\ &= P(\mathsf{Y},\mathsf{S})P(\mathsf{Z}|\mathsf{Y},\mathsf{S})P(\mathsf{X}|\mathsf{Z}) \\ &= P(\mathsf{Y})P(\mathsf{S})P(\mathsf{Z}|\mathsf{Y},\mathsf{S})P(\mathsf{X}|\mathsf{Z}) \end{aligned} \tag{4.2}$$

since Y and S are independent.

- Involuntarily Engaged - Endogenous ($\mathsf{E} \rightarrow \mathsf{Y} \rightarrow \mathsf{Z} \rightarrow \mathsf{X}$):

$$\begin{aligned} P(\mathsf{E},\mathsf{Y},\mathsf{Z},\mathsf{X}) &= P(\mathsf{E})P(\mathsf{Y}|\mathsf{E})P(\mathsf{Z}|\mathsf{Y},\mathsf{E})P(\mathsf{X}|\mathsf{Z},\mathsf{Y},\mathsf{E}) \\ &= P(\mathsf{E})P(\mathsf{Y}|\mathsf{E})P(\mathsf{Z}|\mathsf{Y})P(\mathsf{X}|\mathsf{Z}) \end{aligned} \tag{4.3}$$

- Voluntarily Engaged - Endogenous ($\mathsf{I} \rightarrow \mathsf{Y} \rightarrow \mathsf{Z} \rightarrow \mathsf{X}$):

$$\begin{aligned} P(\mathsf{I},\mathsf{Y},\mathsf{Z},\mathsf{X}) &= P(\mathsf{I})P(\mathsf{Y}|\mathsf{I})P(\mathsf{Z}|\mathsf{Y},\mathsf{I})P(\mathsf{X}|\mathsf{Z},\mathsf{Y},\mathsf{I}) \\ &= P(\mathsf{I})P(\mathsf{Y}|\mathsf{I})P(\mathsf{Z}|\mathsf{Y})P(\mathsf{X}|\mathsf{Z}) \end{aligned} \tag{4.4}$$

4.4 Challenges in BCI decoding

Type	Related with	Causal Shift to	Affects
Stimulus Shift	Experimental Settings	$P(\mathsf{S})$	Involuntarily Engaged and Exogenous
Shifts on Subject's Intention	Experimental Settings	$P(\mathsf{I})$	Voluntarily Engaged and Endogenous
Shifts on Environmental Conditions	Experimental Settings	$P(\mathsf{E})$	Involuntarily Engaged and Endogenous
Acquisition Shift	EEG Data	$P(\mathsf{X}|\mathsf{Z})$	All Cases
Subject Shift	Subjects	$P(\mathsf{Z}|\cdot)$	All Cases
Class Imbalance	Associated Labels	$P(\mathsf{Y})$	Exogenous

Table 4.2: All possible shifts in the joint distribution $P(\mathsf{X},\mathsf{Y},\mathsf{Z},\cdot)$ in brainwave decoding for all causal factorization scenarios described in our causal framework.

Through the above described causal breakdown, we can easily conclude that there are major challenges, that machine learning (statistical) models usually face when dealing with BCI brainwave decoding tasks. It is evident that we are in need of machine learning techniques that do no rely solely on the i.i.d. assumption but they are bound to these causal mechanisms [99].

Data distribution shifts constitute the main factor that hurt generalization of machine learning models. Using the above described causal framework, it is possible to recognise all possible distribution shifts associated with each BCI paradigm and formulate new strategies to mitigate these problems. Let $P_{\mathcal{S}}(\cdot)$ denote a distribution where a shift has been applied to. Table 4.2 summarizes some of the various shifts in brainwave decoding as well as the causal sub-module they usually affect. These distribution shifts can have various sources ranging from the experimental setup, available EEG data to the nature of the brain itself.

BCI researchers need to understand the sources of all possible changes or distribution shifts associated with these core causal variables of interest when designing machine learning models for brainwave analysis. In this section, we outline the core challenges that BCI machine learning models face and we further propose how well-established - as well as newly emerging - techniques can tackle these problems.

4.4.1 Experimental Settings

The experimental conditions (under which a BCI decoding task takes place) are crucial and they need to be designed carefully to avoid any possible distribution shift between the training and test EEG sets. BCI researchers need to take into account the following factors:

1. **Demographics.** Demographics (e.g. sex, age and handedness) are considered to be the cause of differences in brain activation patterns among individuals during the execution of the same task. It is advisable that the subjects' demographics in training EEG set resemble the population represented in the test EEG set. If the goal is a global population-agnostic brainwave decoder, the training EEG set need to be diverse in terms of available subjects in order to capture as large part of the population as possible.

2. **Stimuli Shifts $P_{\mathcal{S}}(\mathsf{S})$.** Brainwave decoders, that rely on the presence of stimuli to modulate the necessary brain activity, are heavily dependent on the exact type and form of this external sensory input. Therefore, the stimuli used for the collection of the training EEG set and the stimuli used during the deployment of the machine learning model need to be perfectly aligned and any possible shifts should be avoided. For example, in the case of a steady state

visually evoked potentials (SSVEP) BCI-speller, the specific frequencies that correspond to each alphabetical letter needs to be the same in both training and test EEG sets.

3. **Shifts on Subject's Intention $P_{\mathcal{S}}(\mathsf{I})$.** Subject's intention is the starting point, and therefore vital, in all voluntarily engaged and endogenous brainwave decoders. It is important though that these intentions are aligned between the training and test EEG sets since different intentions can lead to different BCI-relevant neural activity. For example, in a Motor-Imagery BCI, a subject can imagine raising his/her right hand while another subject can visualize of closing his/her right fist. Although both movement include the movement of the right hand, they can lead to very different brain activities which might mislead the brainwave decoder. It is the BCI researcher's duty to ensure a baseline behaviour for all involved participants during the data collection as well as the deployment phase in order to avoid any unpredictable distribution shifts.

4. **Shifts on Environmental Conditions $P_{\mathcal{S}}(\mathsf{E})$.** The environment factor is essential for involuntarily engaged and endogenous brainwave decoders and therefore, as in the case of stimuli and subject's intention, the environment in the training EEG set should resemble as close as possible the environment represented in the test EEG set.

4.4.2 Training EEG Data

Data Scarcity

One of the main challenges in brainwave modeling is the lack of labeled data. This is because the acquisition process is difficult and time-consuming. Subjects must spend hours wearing an EEG headset and performing consecutive tasks, which can be tiring and uncomfortable, even when dry electrodes are used.

To increase the amount of available training samples, machine learning scientists usually turn to data augmentation. **Data augmentation** is defined as the process to generate artificial samples from the available brainwave data by applying certain shifts (e.g addition of Gaussian noise) to it. This process enriches the available joint distribution $P(X, Y)$ - rather than only $P(X)$ - making it an available candidate for both causal and anti-causal brainwave tasks. Through this technique, the new training set includes data from different distribution shifts, which makes the learnt model more robust, solves partially the issue of data scarcity and leads to better generalization. In the field of BCI, various data augmentation techniques have been introduced [122]: Generative adversarial networks (e.g. [123], [124], [125]), noise addition (e.g. [126], [127]), overlapping sliding windows (e.g. [128]), recombination of segmentation (e.g. [129]) and geometric approaches (e.g. [130]). Data augmentation can lead to an increased performance regardless of the causal direction of the task. For example, [131] demonstrates increased performance in SSVEP classification (voluntarily engaged and exogenous task) while [126] shows improved Motor Imagery classification performance (voluntarily engaged and endogenous task).

Like data augmentation, pre-training's goal is to enrich the training set distribution. Unlike data augmentation though, using **pre-training** does not mean to create or obtain more examples from the joint distribution $P(X,Y)$. This technique is based on the assumption that if a machine learning model is trained on a huge, diverse and close to the targeted brainwave task dataset, it will capture extra information from all of these various distributions, can benefit both causal and anti-causal models and achieve stronger BCI generalization as it is demonstrated in [132] for enhanced intracranial EEG decoding or in [133] for improved Motor-Imagery EEG classification (voluntarily engaged and endogenous task).

Semi-supervised learning aims to leverage huge amount of unlabelled brainwave data in an effort to generate statistical models that achieve better generalization compared to models that are learnt using only a limited amount of available labelled brainwaves. In essence, this technique utilizes large amount of unlabelled data to further enhance $P(X)$. Unlike the previous two techniques, as also described in ([96], [95] and [134]), it is futile when applied to causal tasks ($X \rightarrow Y$) since any additional information, that can improve $P(X)$, will have no influence on $P(Y|X)$ - the quantity that the machine learning model is concerned with. Therefore, this technique should only be used in anti-causal tasks. [135] describes a state-of-the-art BCI system that utilized semi-supervision for epileptic seizure detection (an endogenous - passive task and anti-causal for brainwave decoders). [136] describes an online semi-supervised BCI for P300 speller (voluntarily engaged and exogenous task).

Finally, **Self-Supervised Learning (SSL)** consists of two steps: training on a large unlabelled dataset and fine-tuning on few labelled brainwave samples. As in the case of semi-supervised learning, this technique aims to enrich $P(X)$, a quantity only informative for anti-causal (mostly brainwave decoders) as far as Machine Learning is concerned. In (e.g. [137], [138]) SSL-learned networks consistently outperformed purely supervised deep neural networks in anti-causal paradigms like sleep stage detection (involuntarily engaged and endogenous task).

Acquisition Shift $P_{\mathcal{S}}(\mathsf{X}|\mathsf{Z})$

The data acquisition can be undertaken with various EEG recorders with completely different specifications (e.g. number and type of electrodes or sampling frequency). As a result, this incompatibility of available EEG datasets makes it very difficult to combine them. Although perfectly aligning the EEG recorders after the data collection has taken place is unrealistic, three signal processing steps can be taken in an effort of combining different EEG training sets:

1. Re-sample all training sets to a common sampling frequency. This process should be treated with care so as to avoid aliasing issues and important temporal information loss

2. Retain only the common EEG sensors according to the international 10-20 system [139]

3. Perform a type of normalization (e.g. min-max normalization) to map the values of the EEG signals within the same range while maintaining their original distribution

4. Perform signal re-referencing to a common EEG sensor - if it is vital for a specific brainwave analysis

4.4.3 Subject Shifts

Our brain is a complex system that is not only structured but also functions differently from one person to another, a phenomenon that is often referred to as population shift or inter-subject variability. Each subject has a unique brain anatomy and functionality of each individual that results in a variety of different neural activity patterns that can be observed on the surface of the brain using EEG signals. Additionally, brain functionality does not only differ between subjects but across time as well. According to various studies [140], a certain variability in the acquired EEG signal can be observed when the same subject attends the same experiment in different times (sessions), a phenomenon known as inter-session variability. In the case of BCI brainwave decoders, subject shifts can occur in all possible identified scenarios in our causal framework (i.e. $P_{\mathcal{S}}(\mathsf{Z}|\mathsf{S})$, $P_{\mathcal{S}}(\mathsf{Z}|\mathsf{Y},\mathsf{S})$, $P_{\mathcal{S}}(\mathsf{Z}|\mathsf{Y})$, $P_{\mathcal{S}}(\mathsf{Z}|\mathsf{I})$). Therefore, we should turn our attention to domain-specific (BCI-specific) methods that are specifically designed to tackle these issues, found only in the area of brainwave analysis.

Over the last few years, there has been an increasing interest towards this research direction with many strong empirical evidence of increasing the performance of various BCI models. [141]

projects data into a subject-invariant space while [142] uses an adversarial inference network that learns subject invariant features. In an effort to design enhanced subject-independent machine learning models, [18] describes a novel technique to explicitly align feature distributions at various layers of the deep learning model. By utilizing both statistical and trainable methods (inspired by covariance-based alignment methods used in Riemannian geometry [143]). In Chapter 6, we introduce a dynamic framework that approximates the subject distribution shift in the true BCI task-related brain activity among different subjects by using dynamic convolutions based on an input-dependent attention vector.

Artifacts

The acquired EEG signal is extremely vulnerable to undesired noise, which can result in various artifacts [144]. As artifacts can be defined abnormal signals which can contaminate the normal brainwaves and severely interfere with machine learning-based brainwave decoders, since they can alter the measured EEG signal. Therefore, these artifacts need to removed from the EEG sets used with machine learning decoders [145]. EEG-related artifacts can be categorized to extrinsic (environmental artifacts and artifacts due to experimental errors) and intrinsic or physiological artifacts (e.g. eye blinks, muscle activity, heartbeat) [146]. The environmental artifacts can be eliminated by a simple filtering [147] while the proper experimental procedure can reduce or eliminate the experimental errors [145]. Finally, the elimination of intrinsic artifacts requires specialized techniques which include Regression [148], Independent Component Analysis (ICA) [149], Empirical-Mode Decomposition (EMD) [150], Blind Source Separation (BSS) [151] and Wavelet Transform [152].

4.4.4 Associated Task Labels

It is important to make sure that the training set for a machine learning model is as similar as possible to the deployment set, in terms of the distribution of classes (task class balance). This is because if the training set is imbalanced, the model may not be able to generalize well to new unseen data during deployment.

In the cases where this is not possible, specific techniques should be employed to correct the bias in estimating the training loss. For instance, [153] uses a BCI method for seizure detection that effectively moderates the bias in performance caused by imbalanced class distribution by assigning different weights to the EEG samples.

Step-by-step guidelines for designing BCI decoders using our causal framework

Given an EEG dataset $\mathcal{D}$ (or multiple datasets):

1. Ensure that all experimental conditions (stimuli, environmental conditions, demographics and subject's intentions) are designed carefully and resemble the deployment environment.

2. Gather all necessary and important information about $\mathcal{D}$. If needed, remove all artifacts. If you combine multiple datasets, then make sure to align the sampling frequency, re-reference and normalize the EEG signals.

3. Using the introduced causal framework, identify the causal scenario of the brainwave decoding model.

4. If $\mathcal{D}$ is sparse, then deploy data augmentation or pre-training techniques for both causal and anti-causal tasks. Or self-supervised and semi-supervised learning for anti-causal tasks only.

5. Using the causal factorization of the targeted BCI decoding case, identify the possible distribution shifts and employ appropriate solutions.

4.5 Discussion

Causal Reasoning enables us to identify possible distribution shifts that can hurt the generalization of machine learning-based BCI brainwave decoders. In the previous section, we explore how general machine learning practices as well as specifically designed techniques can contribute into solving many of the identified causal challenges. Although these techniques are crucial by themselves, we need to identify appropriate evaluation schemes in order to measure and compare their performance and develop plans on making improvements. So far, in brainwave modeling for BCIs, the performance of trained models is measured based on some unseen parts of the available EEG dataset. Splitting brainwave datasets into training and testing parts comes with certain limitations and inherited biases. Therefore, EEG datasets need to be diverse in terms of the available subjects (to capture a large part of the population) as well as include multiple sessions per subject. In that way, various machine learning models could be tested (with the appropriate splitting) for their performance across various tasks as well as their robustness in a number of the above identified challenges and data distribution shifts like inter-subject and inter-session variability.

The machine learning community of BCIs is in need of appropriate benchmarks, similar to well-established efforts in other areas like Computer Vision (e.g. [154], [155]), that can capture multiple scenarios and act as a valid indicator of the performance of different brainwave decoders. On that front, the Benchmarks for EEG Transfer Learning (BEETL) competition [17] took place at the top-tier NeurIPS (Neural Information Processing Systems) 2021 conference. This benchmark provides a referencing platform for transfer learning strategies and performances on two widely-used BCI paradigms (motor-imagery and sleep stage detection). Consisting of two tasks, this benchmark includes two challenges (1. transfer across subjects and 2. transfer across subjects and data sets) and promotes the field of brainwave analysis towards the use of big data. Although this competition was a significant milestone, more similar initiatives are needed in that direction that will capture more BCI paradigms and distribution shift scenarios.

4.6 Conclusions

In this chapter, we take a Causal Reasoning approach to analyze the task of brainwave modeling. We introduce a causal framework that relies on two core properties (the presence of task stimulus and the voluntary engagement of the subject) and can break down every BCI paradigm to its independent components. To the best of our knowledge, this the first study to combine machine learning and Causal Reasoning in the field of BCIs and presents a unified causal framework. This is a theoretical study, but we believe that this causal framework represents a crucial advancement, poised to serve as a guiding principle for the development of advanced BCI brainwave decoders.

In the following two chapters, this causal framework will guide us to identify potential problems in the MI classification problem using machine learning models. As described above, Motor-Imagery is an endogenous (the brain activity is inherently modulated, without the intervention of any external stimulus) and voluntary engaged task (the subject willingly generates a particular brain activation pattern) with the following general causal diagram:

$$\mathsf{X} \leftarrow \mathsf{Z} \leftarrow \mathsf{Y} \leftarrow \mathsf{I} \tag{4.5}$$

In the next chapters, we perform within-dataset experiments, using either in-house collected or publicly available MI EEG datasets, meaning that a baseline behaviour for all involved participants can be guaranteed and any unpredictable distribution shift on subject's intention $P_{\mathcal{S}}(\mathsf{I})$ is avoided.

Therefore, for the remainder of the Thesis, we will consider that I and Y are aligned and we are interested in the following causal diagram:

$$\mathsf{X} \leftarrow \mathsf{Z} \leftarrow \mathsf{Y} \tag{4.6}$$

In a MI classification problem, we want to accurately predict the mentally performed task from a recorded EEG signal. Mathematically, given an input EEG signal X, we train a statistical model to predict the correct MI task Y, which can be the imagery movement e.g. of a hand or foot. In essence, this statistical model tries to estimate the conditional probability $P(Y|X)$ using an appropriate objective function.

Using the above basis, we can define an MI EEG classification task as an anti-causal problem, since the true MI intention (observed with the MI label Y) can be considered the cause of the recorded EEG signal X. As a consequence of the above anti-causal definition and causal diagram, we can explore the problem of MI EEG classification through the following causal factorization:

$$P(\mathsf{X}, \mathsf{Y}, \mathsf{Z}) = P(\mathsf{X}|Z)P(\mathsf{Z}|\mathsf{Y})P(\mathsf{Y}) \tag{4.7}$$

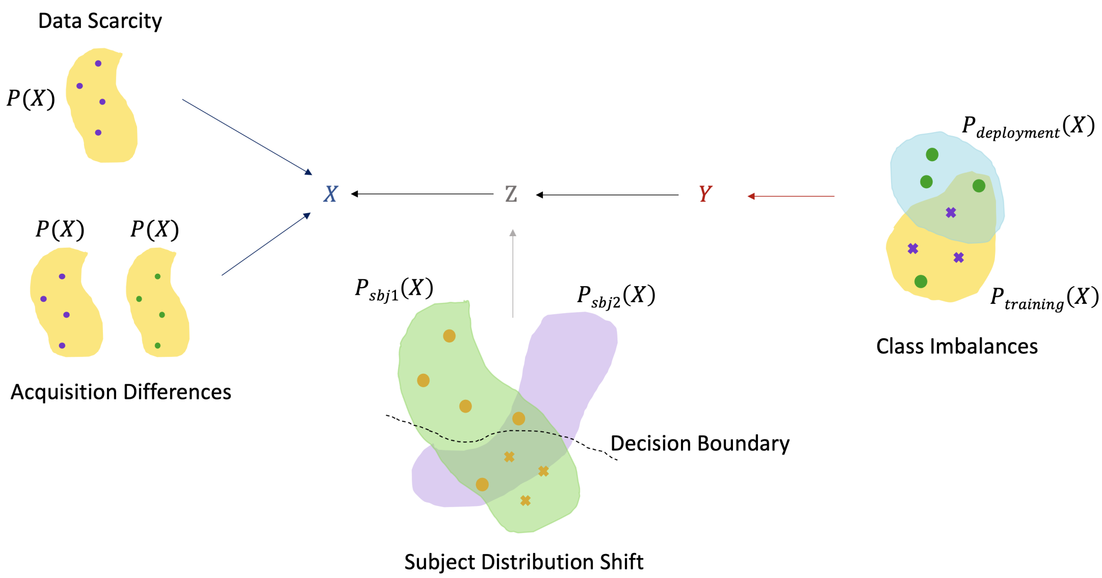

Figure 4.4: Key challenges in Machine Learning for a MI EEG classification task. X represents input EEG signals, Z the true unobserved brain activity, Y the associated MI labels. • and × represent EEG signals of different labels. Dots represent data points of any label and their color represent different EEG acquisition devices.

5
BrainWave Scattering Net

Contents

In this chapter, we introduce BrainWave-Scattering Net, a lightweight deep neural network architecture that is based on the joint time-frequency wavelet scattering transform, which was published in **K. Barmpas**, Y. Panagakis, D. Adamos, N. Laskaris, and S. Zafeiriou, *"BrainWave-Scattering Net: A lightweight network for EEG-based motor imagery recognition"*, in Journal of Neural Engineering (2023). Moreover, we introduce a subject selection / subject pooling framework that uses the above method and was published in **K. Barmpas**, Y. Panagakis, D. Adamos, N. Laskaris, and S. Zafeiriou. *"Subject Selection Framework to Improve Personalised Models for Motor-Imagery BCIs via Wavelets and Graph Diffusion"* in ICLR 2024 Workshop on Learning from Time Series For Health (2024).

5.1 Introduction

Guided by the causal framework of Chapter 4 and using the causal factorization for the MI classification problem (Figure 4.4), in this chapter we will focus mainly on the problems related to the training data X.

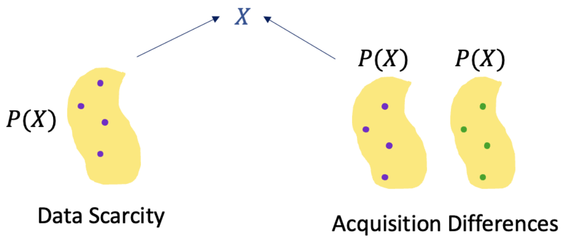

Figure 5.1: Key challenges in Machine Learning for a MI EEG classification task regarding the training EEG signals X. X represents input EEG signals. Dots represent data points of any label and their color represent different EEG acquisition devices - Part of Figure 4.4.

Challenge 1: One of the renowned challenges in motor-imagery classification problem - as in any medical-related machine learning problem - is the scarcity of labelled data due to the lengthy acquisition process (e.g. [156],[157],[158]). Subjects are required to spend hours in a laboratory facility performing successive motor-imagery tasks [159]. This process has been reported to cause fatigue and discomfort, even when devices with dry electrodes are utilized.

Challenge 2: To make things worse, due to the wide variety of available EEG recorders in the market, the data acquisition can be undertaken with various devices which have completely different specifications (e.g. number of electrodes, sampling frequency to name just a few), making the combination of available EEG datasets extremely difficult [17].

For many years, MI-BCI decoders used to rely on manual feature extraction, a process which usually requires significant neurological expertise and experience in EEG signal analysis. The massive progress in the field of Deep Learning has largely alleviated the need for manual feature extraction. While various deep BCI neural network architectures have achieved state-of-the-art performance (e.g. [62], [63], [64]), their conception as "black-boxes" usually means an incomprehension of the underlying extracted features. In addition, the required large amount of data to train these networks usually prohibits the training of high performed personalised BCI models where only a handful of data is available. Designing personalised BCI models effectively solves the issue of inter-subject variability [160] - data distribution shift across subjects due to differences in brain anatomy and function between individuals (e.g. [161], [162]) - that general BCI models need to tackle. But the scarcity of labelled EEG data prohibits the training of deep BCI personalised networks leading often to overfitting (**Challenge 1**). Lastly, MI-BCI models trained on a specific number of electrodes struggle to generalize to other electrode configurations due to the intricate relationship between electrode placement and neural signal representation. The spatial distribution of electrodes profoundly influences captured neural activity patterns during MI tasks, making decoders finely attuned to specific configurations (**Challenge 2**).

In this chapter, to address the above mentioned challenges, we introduce the BrainWave-Scattering Net, a lightweight network architecture that is based on a novel, fully-learnable architecture inspired by the joint time-frequency scattering transform [163]. By utilizing a trainable wavelet-based scattering transform pipeline, BrainWave-Scattering Net is able to learn representations that capture the important underlying characteristics and properties of an EEG signal for MI classification tasks. Indeed, by conducting experiments in two different publicly available MI datasets and one collected in-house MI dataset, it is demonstrated that BrainWave-Scattering Net can achieve high performance compared to other state-of-the-art deep architectures, namely DeepConvNet [62], ShallowConvNet [62], EEGNet [63] and EEG-Inception [64], while having considerably less trainable parameters. In addition, its lightweight nature allows considerably faster training time as well as training of efficient personalised BCI models with limited amount of given data (**Challenge 1**). Furthermore, it demonstrates better explainability properties and device transferability properties (**Challenge 2**).

Finally, in this chapter, a novel subject selection framework is introduced that blends ideas from discriminative learning (based on continuous wavelet transform) and graph-signal processing (over the sensor array). This selection framework augments the data available to craft a personalised model (**Challenge 1**), by incorporating additional signals from subjects with similar brain activations. Through experimentation with the BrainWave-Scattering Net and a publicly available MI dataset, we showcase enhanced personalised performance for MI-BCIs. Notably, it proves particularly advantageous for subjects who initially demonstrated sub-optimal personalised performance.

In summary, this chapter's contributions are as follows:

- A lightweight deep neural network architecture is introduced, based on the joint time-frequency wavelet scattering transform initially described in [163]. The novelty lies in the adaptation of the transform to handle EEG data and the introduction of learnable wavelet filters that outperform the fixed wavelets used in the original method.
- Similar - and in some cases better - performance is demonstrated compared to other state-of-the-art deep BCI architectures, with considerably fewer trainable parameters as well as shorter training time.
- Better interpretability insights into properties of brain signals in the field of Motor-Imagery are provided compared to "black-box" approaches present in other BCI deep learning networks.
- Leading transferability to lower density EEG sensor arrays is provided compared to other state-of-the-art deep BCI architectures without the need for model re-training. In addition, high levels of robustness during inference are demonstrated when some of the input EEG signals are tampered with.
- A subject selection framework, harnessing wavelet transform and graph diffusion distances, is introduced achieving increased personalised performances of subjects who initially exhibited poor performance.

The remainder of this chapter is organized as follows: Section 5.2 describes related work in the literature of Differentiable Signal Processing and Section 5.3 outlines the network's architecture. Section 5.4 consists of the experimental part, where performance results and comparisons on different datasets are detailedly presented. It also contains an analysis of the main advantages of the network as well as a discussion part to demonstrate the importance of the trainable element in our network by comparing and contrasting standard wavelet techniques with their equivalent data-driven approaches. Section 5.5 outlines our subject selection framework along with the corresponding experimental results. The last section 5.6 summarizes and concludes the work presented in this chapter.

The main objectives in this chapter are:

to introduce a novel deep architecture and a subject selection / subject pooling framework that face the challenges of **data scarcity** and **acquisition differences** in the MI-BCIs

5.2 Related Work

Although deep architectures achieve state-of-the-art performance across different tasks, they usually tend to be comprised of thousands of trainable parameters. In addition, their extracted learned features are not interpretable due to their unconstrained nature during training. Therefore, several neural network architectures in the field of Differentiable Signal Processing have replaced standard convolutional layers with structured trainable filters exploiting the characteristics of known filter families, such as Gabor filters ([164], [165] and [166]), Sinc [167] filters or Spline filters [168]. In this way, their gain is double: reduced model complexity and improved interpretable layers. In the field of audio processing, SincNet [167] uses sinc bandpass filters as convolutional kernels (w) which are described by the following equation:

$$w(t) = 2f_2 \ \mathrm{sinc}(2\pi f_2 t) - 2f_1 \ \mathrm{sinc}(2\pi f_1 t) \tag{5.1}$$

where $sinc(x) = \frac{sin(x)}{x}$. By having the frequencies f_1 (low cutoff frequency) and f_2 (high cutoff frequency) to be the only trainable parameters rather than the whole function provides enhanced interpretability properties in the network. [165] approximates the mel-filterbank in the time-domain using the first order scattering transform as:

$$Mx(t,n) \approx |x * w_n|^2 * |\varphi|^2(t) \tag{5.2}$$

where $Mx(t,n)$ denotes mel-filterbank consisting of n triangular filters, φ denotes the Hanning window, and w_n the n-th Gabor wavelet described by:

$$w_n(t) \approx e^{-2\pi i \eta_n t} \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{t^2}{2\sigma_n^2}} \tag{5.3}$$

where η_n denotes the desired center frequency and σ_n the width of the n-th filter. Since Gabor wavelet is defined as a complex function, in order to define a complex-valued convolution with N filters, [165] uses an intermediate step which computes 2N real-valued filters, corresponding to the real and imaginary parts, and performs a series of pooling and stride operation along the channel axis to obtain the squared modulus using the adjacent pair filters. Although [165] uses initiallized filters that approximate mel-filterbanks, they in turn are jointly freely tuned with the remaining convolutional architecture. Aiming to increase interpretability, in a follow-up work [164] trainable Gabor filters are used as convolutional kernels (w) which are represented by the following equation:

$$w(t) = e^{2\pi i \eta_n t} \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{t^2}{2\sigma_n^2}} \tag{5.4}$$

The trainable parameters, here, are only the center frequency η_n and the bandwidth $\frac{1}{\sigma_n}$ (both expressed in normalized frequency units in [-$\frac{1}{2}$, +$\frac{1}{2}$]) as opposed to the full Gabor filter function.

5.3 BrainWaveScattering Net

5.3.1 Preliminaries

When deriving features from an input signal $x(t)$ for classification purposes like in a MI-BCI, these representations need to be stable under small translations and deformations. Wavelet scattering transform as described in [169] builds these representations through a non-linear, unitary transform, which delocalizes signal information into scattering decomposition paths. They are computed with a cascade of wavelet transforms [170] followed by a modulus operator and a filtering on the time domain at the end to ensure local invariance to translations [169].

Let $*$ denote the convolution operator, mathematically the wavelet scattering transform Sx of a signal $x(t)$ along time with a cascade of wavelets $(\psi_{\lambda_1}(t), \psi_{\lambda_2}(t), ..., \psi_{\lambda_n}(t))$ and an average filtering on the time domain $\phi(t)$ is defined as [169]:

$$Sx = |||x(t) * \psi_{\lambda_1}(t)| * \psi_{\lambda_2}(t)| ... * \psi_{\lambda_n}(t)| * \phi(t) \tag{5.5}$$

In many tasks, only the first $(Sx_1 = |x(t) * \psi_{\lambda_1}(t)| * \phi(t))$ and second order $(Sx_2 = ||x(t) * \psi_{\lambda_1}(t)| * \psi_{\lambda_2}(t)| * \phi(t))$ wavelet scattering coefficients are utilized since higher order coefficients tend to have negligible energy [171]. Therefore, the first order scalogram $|x(t) * \psi_{\lambda_1}(t)|$ is significant since it describes the time-frequency intensity of $x(t)$ by capturing the Hilbert-demodulated temporal envelope of each frequency sub-band [169].

While the time wavelet scattering transform of a signal $x(t)$ successfully describes temporal modulations [169], it fails to capture more sophisticated time-frequency features. For instance, the effects of frequency-dependent time-shifts are not captured by time wavelet scattering transform. Spectro-temporal relationships can be beneficial in the MI-decoding since brainwaves of, at least, two distinct brain rhythms are known to involve in MI-execution, and at different instances. Specifically, the alpha-band desynchronization (known as ERD) marks the initiation of the MI-movement and beta-band rebound its termination (known as ERS) [33]. [163] proposes the joint time-frequency wavelet scattering transform which brings the two-dimensional filter analysis into the framework of wavelet scattering in order to inherit its invariance and stability properties. The joint time-frequency wavelet scattering transform consists of a first order time scattering transform (using wavelet $\psi_{\lambda_1}(t)$) followed by a two-dimensional wavelet analysis being carried out independently in the time and frequency domain with two one-dimensional wavelets.

$$Sx = ||x(t) * \psi_{\lambda_1}(t)| * \Psi(t, \lambda)| * \phi(t) \tag{5.6}$$

where $\Psi(t, \lambda) = \psi(t)\psi(\lambda)$ is the product of two one-dimensional wavelets functions in time and frequency. Equation 5.6 captures the joint variability of $|x(t) * \psi_\lambda(t)|$ at frequency and time, while the modulus and time-averaging operation $\phi(t)$ ensures time-shift invariance and time-warping stability ([163], [172], [173]).

A common choice for the mother wavelet $\psi_\lambda(t)$ in the wavelet scattering transform Sx is the Gabor wavelet due its property to minimize the product of its standard deviations in the time and frequency domain [174]. This constitutes the Gabor envelope the best time-frequency localization window according to the Heisenberg uncertainty principle [174]. The complex Gabor 1-D wavelet is defined as:

$$\psi_\lambda(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} e^{2\pi i \lambda t} \tag{5.7}$$

where $\frac{1}{\sigma_n}$ denotes the bandwidth and λ the center frequency.

5.3.2 Architecture

Here, BrainWave-Scattering Net (Figure 5.2), a lightweight deep architecture based on a novel learnable joint time-frequency scattering transform for EEG-based motor-imagery recognition, is introduced. Its architecture consists of four main blocks: The first two blocks extract frequency and spatio-temporal information from the EEG signals through a cascade of learnable wavelet transforms. The third block performs a time-averaging operation to ensure shift-invariance while the last block performs the spatial analysis of the signals. In this section, these four building blocks (Table 5.1) are detailedly described and are presented using tensor operations. Without loss of generality, the batch size is omitted in the following formulation as well as the stride and padding operations in the following mathematical formulas.

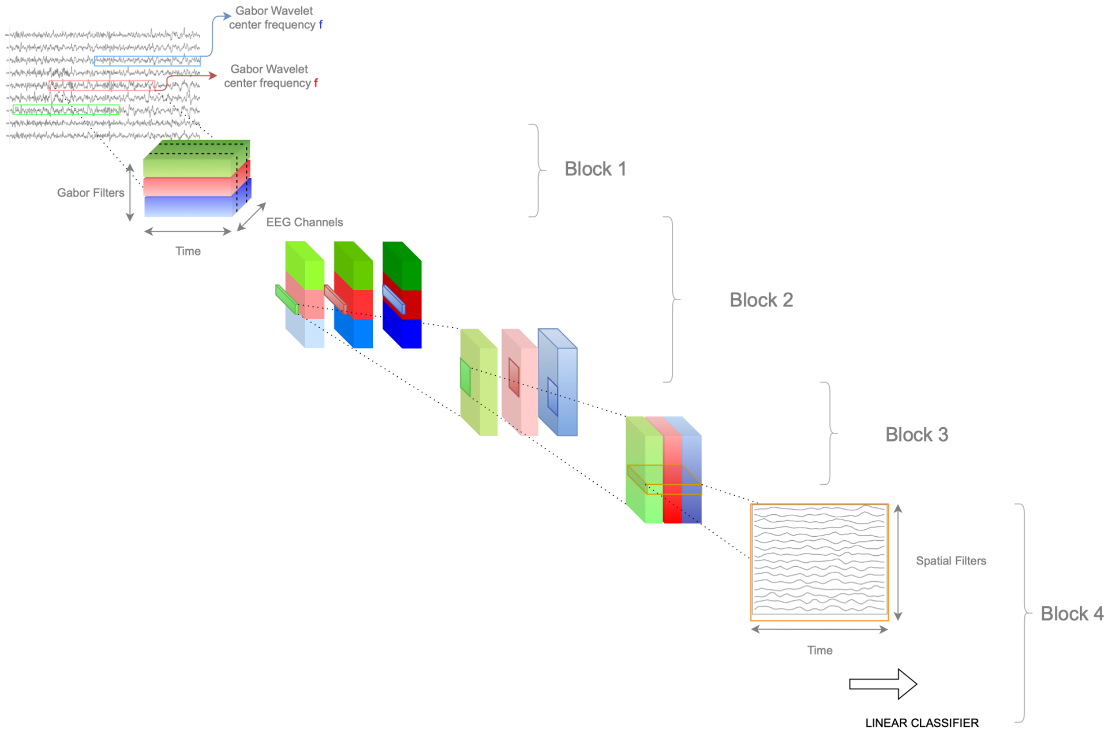

Figure 5.2: Visualization of the BrainWave-Scattering Net architecture. In Block 1, different colors represent extracted scalograms from different Gabor wavelets. In Block 2, each "slice" represents a different EEG channel with all its extracted scalograms stacked and ordered based on the corresponding wavelet's frequency. In Block 2, depthwise convolutions, first across time and then across frequency, are denoted with different colors. The depthwise convolution across channels implements the spatial filtering and is denoted with orange color. Finally, a linear classifier uses the features of the spatial analysis to classify the MI task.

Block 1 - Extraction of Frequency Information: The first block of the BrainWave-Scattering network computes the 1st order scalogram of the joint time-frequency scattering transform. Mathematically, let $\mathbf{x}(t) \in \mathbb{R}^T$ denote a one-dimensional input EEG signal, where T is the number of initial EEG time points, and $\boldsymbol{\psi}_\lambda(t)$ be a wavelet. This block computes the 1st order scalogram $\mathbf{X}(\lambda, t)$, the amplitude of the wavelet transform, for each EEG channel separately:

$$\mathbf{X}(\lambda, t) = |\mathbf{x}(t) * \boldsymbol{\psi}_\lambda(t)| \tag{5.8}$$

where t denotes the time and λ the frequency. To perform this operation, the raw input signal from each EEG channel is convolved with a wavelet kernel with size $(1, W) = (1, \frac{F_s}{2})$ where F_s is the sampling frequency. Unlike standard convolution layers where all the parameters are fully trainable, this layer follows a particular wavelet function. More specifically, it follows the Gabor wavelet equation, since Gabor filters are optimally localized in time and frequency [175] and can capture valuable features from the EEG signals. In this layer only the real part[13] of the wavelet is used to take advantage of its symmetric properties. In other words:

$$\boldsymbol{\psi}_{\lambda}(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} cos(2\pi\lambda t) \tag{5.9}$$

where t $\in [-\frac{W}{2}, ..., \frac{W}{2}]$ and $\frac{1}{\sigma}$ denotes the bandwidth and λ the normalized frequency of the Gabor wavelet and these two properties are the only trainable parameters of this layer. Although, $\lambda = \frac{F}{F_s}$, where F denotes the actual frequency and F_s is the sampling frequency both measured in Hz, is a trainable parameter, it needs to be restricted to satisfy the Nyquist Theorem [176], therefore:

$$0 \leq F \leq \frac{F_s}{2} \rightarrow 0 \leq \lambda \leq \frac{1}{2} \tag{5.10}$$

The initialization step of this layer is crucial in order to properly guide the network to extract valuable frequency information for the targeted motor-imagery task. As described in [33], in Motor-Imagery two frequency bands carry the most information: alpha band [8, 13Hz] and beta band [13, 30Hz]. Certain subjects present also some activity in the lower gamma band [30, 40Hz]. Therefore, the different λ values, during initialization, are evenly spaced in the following range:

$$\frac{8}{F_s} \leq \lambda \leq \frac{40}{F_s}(Hz) \tag{5.11}$$

The $\frac{1}{\sigma}$ bandwidth is initialized with the same value for all wavelet filters. The "right" choice of its initial value depends on different datasets and should be treated during the hyper-parameter tuning phase of the network. Nevertheless, as described in [164], a reasonable range for the σ bandwidth values is $[4\sqrt{2log2}, 2W\sqrt{2log2}]$ "such that the full-width at half-maximum of the frequency response is within 1/W and 1/2" (we refer the reader to an ablation study in section 5.4.8 that showcases how the choices of the bandwidth value can influence the network's performance).

The input to this block, and inherently the input to our network, can be described as a 2nd-order tensor $\mathbf{I}$ which consists of C (where C is the number of EEG channels) row-vectors $\mathbf{x}(t) \in \mathbb{R}^T$. More specifically in this layer, F one-dimensional Gabor filters are convolved with each input EEG

[13]It is straightforward to use the Gabor wavelet directly by doubling the number of output filters as described in [164], but this leads to computational and coding overhead

row-vector $\mathbf{x}(t)$ to compute the 1st order scalogram using trainable Gabor wavelets as convolutional kernels. The result of each one of these C one-dimensional convolutions is a matrix $\mathbf{X}(\lambda, t) \in \mathbb{R}^{F \times T}$ where the wavelet filters are ordered based on their normalized frequencies. By taking the modulus and stacking these C matrices together, the 1st order scalogram of all EEG channels is produced in a three-dimensional tensor format $X(c, \lambda, t) \in \mathbb{R}^{C \times F \times T}$.

Finally, the computation of the scalogram tensor $X(c, \lambda, t) \in \mathbb{R}^{C \times F \times T}$ is followed by an average pooling with size of 4 across time to reduce the sampling rate of the signal by 4.

Block 2 - Extraction of Spectro-Temporal Information: The second block of the BrainWave-Scattering network computes the joint spatio-temporal features of the joint-scattering transform. Mathematically, let $\mathbf{X}(\lambda, t) \in \mathbb{R}^{F \times \frac{T}{4}}$ be the 1st order scalogram computed in Block 1 for one EEG channel after the pooling operation, this block performs the following operation:

$$\mathbf{F}(\lambda, t) = |\mathbf{X}(\lambda, t) * \mathbf{\Psi}(\lambda, t)| \tag{5.12}$$

where $\mathbf{\Psi}(\lambda, t) = \boldsymbol{\psi}(\lambda)\boldsymbol{\psi}(t)^T$ is the product of two one-dimensional functions in time and frequency with dimensions $\boldsymbol{\psi}(t) \in 1 \times \mathbb{R}^{\frac{F_s}{8}}$ and $\boldsymbol{\psi}(\lambda) \in 1 \times \mathbb{R}^{\frac{F_s}{8}}$ since the sampling rate of the signal is reduced by 4 after the average pooling in Block 1. With joint time-frequency scattering our intention is to scatter the information further across each EEG channel independently, to perform the $\mathbf{\Psi}(\lambda, t)$ operation, depthwise convolutions are utilized to explicitly decouple the relationship within and across the different EEG channels, as described in [177]. Using depthwise operations, first across time and then across frequency, this block extracts features for each EEG channel separately capturing useful spatio-temporal relationships. In this layer, the kernel's parameters are fully trainable.

To mathematically describe these operations, we will use the fact that separable convolutions can be obtained by regular convolutions after the application of kernel Canonical Polyadic (CP) tensor decomposition as described in [177]. Therefore, the three-dimensional (3D) tensor $X(c, \lambda, t) \in \mathbb{R}^{C \times F \times \frac{T}{4}}$ can be described as C two-dimensional matrices $\mathbf{X}(\lambda, t) \in \mathbb{R}^{F \times \frac{T}{4}}$ and each kernel of this block operates into these C matrices $\mathbf{X}(\lambda, t)$ separately to compute:

$$\mathbf{F}(\lambda, t) = |\mathbf{X} * \mathbf{\Psi}(\lambda, t)| = |\mathbf{X} * \left(\color{red}{\psi(\lambda)}\color{blue}{\psi(t)^T}\right)| \tag{5.13}$$

Let $\mathbf{\Psi}(\lambda, t) \in \mathbb{R}^{\frac{F_s}{8} \times \frac{F_s}{8}}$ be convolved with $\mathbf{X}(\lambda, t) \in \mathbb{R}^{F \times \frac{T}{4}}$, the resulting feature map $\mathbf{F}(\lambda, t) \in$

$\mathbb{R}^{F\times\frac{T}{4}}$ is defined as:

$$\mathbf{F}(\lambda,t) = |\mathbf{X} * \mathbf{\Psi}| = \Big|\sum_{j=1}^{\frac{F_s}{8}}\sum_{i=1}^{\frac{F_s}{8}}\mathbf{\Psi}(j,i)\mathbf{X}(j+\lambda,i+t)\Big| \tag{5.14}$$

As in [177], applying a low rank-R CP-Decomposition to the kernel tensor $\mathbf{\Psi}$ can be re-written as:

$$\mathbf{\Psi}(\lambda,t) = \mathbf{u}_\lambda \otimes \mathbf{u}_t = \mathbf{u}_\lambda {\mathbf{u}_t}^T \tag{5.15}$$

By combining the last two equations we can write:

$$\mathbf{F}(\lambda,t) = \Big|\mathbf{X} * \mathbf{\Psi}\Big| = \Big|\underbrace{\sum_{j=1}^{\frac{F_s}{8}}\mathbf{u}_{\lambda j}\underbrace{\Big(\sum_{i=1}^{\frac{F_s}{8}}{\mathbf{u}_{t}}_i^T\mathbf{X}(j+\lambda,i+t)\Big)}_{\text{Depthwise Convolution}}}_{\text{Depthwise Convolution}}\Big|$$

$$\mathbf{F}(\lambda,t) = \Big|\mathbf{X} * \mathbf{\Psi}\Big| = \Big|\underbrace{\sum_{j=1}^{\frac{F_s}{8}}\mathbf{u}_{\lambda j}\underbrace{\Big(\sum_{i=1}^{\frac{F_s}{8}}{\mathbf{u}_{t}}_i^T\mathbf{X}(j+\lambda,i+t)\Big)}_{\text{Depthwise Convolution}}}_{\text{Depthwise Convolution}}\Big|$$

Finally, the computation of $F(\lambda,t) \in \mathbb{R}^{F\times\frac{T}{4}}$ is followed by an average pooling with size of 2 across time to reduce the sampling rate of the signal by 2.

Block 3 - Time-Averaging: The third block of the BrainWave-Scattering performs the temporal filtering operation of the joint-scattering transform. More specifically, this block constitutes the operator $\boldsymbol{\phi}(t)$ in the following equation:

$$\mathbf{S}(\lambda,t) = \mathbf{F}(\lambda,t) * \boldsymbol{\phi}(t) = |\mathbf{X}(\lambda,t) * \Big(\boldsymbol{\psi}(\lambda)\boldsymbol{\psi}(t)^T\Big)| * \boldsymbol{\phi}(t) \tag{5.16}$$

To perform this operation, it utilizes depthwise convolutions with stride of 2 and fully-trainable kernel of size of $(1, \frac{F_s}{16})$ since the sampling rate of the signal is totally reduced by 8. In essence, this operation performs a temporal filtering for each EEG channel separately. Therefore, the combined Block 2 and Block 3 feature map $\mathbf{S}(\lambda,t) \in \mathbb{R}^{F\times\frac{T}{16}}$ can be described with tensor operations as:

Block	Layer	Operation	Size	Output	Options
1	Input			(C, T)	
	Reshape			(1, C, T)	
	Gabor Wavelet	$\mathbf{X}(\lambda, t) = \lvert \mathbf{x}(t) * \psi_{\lambda}(t) \rvert$	$(1, \frac{F_s}{2})$	(F, C, T)	Modulus Operator
	Average Pooling		(1,4)	$(F, C, \frac{T}{4})$	
	Dropout				0.25
	Reshape			$(C, F, \frac{T}{4})$	
2	Depthwise Convolution	$\mathbf{F}(\lambda, t) = \lvert \mathbf{X}(\lambda, t) * \mathbf{\Psi}(\lambda, t) \rvert$	$(1, \frac{F_s}{8})$	$(C, F, \frac{T}{4})$	
	Depthwise Convolution		$(\frac{F_s}{8}, 1)$	$(C, F, \frac{T}{4})$	Modulus Operator
	Average Pooling		(1,2)	$(C, F, \frac{T}{8})$	
	Dropout				0.25
3	Depthwise Convolution	$\mathbf{S}(\lambda, t) = \mathbf{F}(\lambda, t) * \boldsymbol{\phi}(t)$	$(1, \frac{F_s}{16})$	$(C, F, \frac{T}{16})$	stride = 2
	Dropout				0.25
	Reshape			$(F, C, \frac{T}{16})$	
4	Depthwise Convolution	Spatial Analysis	(C,1)	$(S, 1, \frac{T}{16})$	ELU Activation - $\lvert\lvert w \rvert\rvert^2 < 1$
	Dropout				0.25
	Flatten			$S \times \frac{T}{16}$	
	Linear Classifier			N	

Table 5.1: Detailed breakdown of each block of BrainWave-Scattering Net architecture where C = number of channels, T = number of time points, F = number of wavelet filters, S = number of spatial filters, and N = number of output classes.

$$\mathbf{S}(\lambda, t) = \mathbf{F}(\lambda, t) * \boldsymbol{\phi}(t) = \sum_{i'=1}^{\frac{F_s}{16}} \boldsymbol{\phi}(i') \mathbf{F}(1+\lambda, i' + t) = \underbrace{\sum_{i'=1}^{\frac{F_s}{16}} \boldsymbol{\phi}(i') \left| \underbrace{\sum_{j=1}^{\frac{F_s}{8}} \mathbf{u}_{\lambda j} \left(\underbrace{\sum_{i=1}^{\frac{F_s}{8}} \mathbf{u}_{t i}^{T} \mathbf{X}(j + \lambda + 1, i + i' + t)}_{\text{Depthwise Convolution - Time}} \right)}_{\text{Depthwise Convolution - Frequency}} \right|}_{\text{Depthwise Convolution - Time-Averaging}} \tag{5.17}$$

Block 4 - Classification: Inspired by Common Spatial Pattern (CSP) [178], this block performs the spatial analysis of the signal. By utilizing a depthwise convolution with kernel of size (C, 1), it extracts the spatial filters of the joint time-frequency scattering transform. Depthwise convolution is utilized to avoid mixure of information across different joint time-frequency scattering feature maps. Similarly to EEGNet [63], each spatial filter is regularized by using a maximum norm constraint of 1 on its weights; $||w||^2 < 1$. Finally, the extracted feature maps are passed into a linear classifier for the BCI motor-imagery classification task.

5.4 Experiments

In the experimental part of this chapter, two publicly available MI datasets are used, namely **PhysioNet** [66] and **BCI IVa Dataset of the 3rd BCI competition** [67], as they were described in Chapter 2 and an in-house collected MI dataset, namely **MyBrainCommands - Standard MI Paradigm** as described in Chapter 3.

5.4.1 Generic Models of Motor-Imagery Recognition

We commence the results sections with presenting the achieved performance when our network was employed to derive generic models of MI-recognition. In this case, the models or decoders, once trained on a large sample of subjects' data, are expected to perform sufficiently to brainwave recordings of unseen subjects. Hence, the validation of their performance had to be cast accordingly, i.e. in a "leave-M-subjects-out" manner. Specifically, 10-fold subject-wise cross-validation was performed for PhysioNet data and MyBrainCommands - Lab Setting data and 5-fold for BCI Dataset IVa. For comparison purposes, the performance achieved by the four state-of-the-art deep learning architectures (described in Chapter 2) is also included, validated in the same way with our models.

In the case of DeepConvNet [62], ShallowConvNet [62], EEGNet [63] and EEG-Inception [64], the models were trained using the Adam optimizer, with batch size of 64, minimizing the cross-entropy loss function and they were run for 150 training iterations. Since our network is lightweight and has the trainable Gabor filters in its first layer, different training settings were used to ensure stability during training, to avoid overfitting and to allow the trainable Gabor filters to quickly adapt to the training data:

- PhysioNet: Adam optimizer was used with learning rate of 0.01 for the first 30 epochs and 0.0001 for the remainder 20 epochs.
- BCI Dataset: Adam optimizer was used with learning rate of 0.01 for the first 50 epochs, 0.001 for epochs between 50-80 and 0.0001 for the remainder 20 epochs.
- MyBrainCommands - Lab: Adam optimizer was used with learning rate of 0.01 for the first 30 epochs and 0.001 for the remainder 120 epochs.

The obtained experimental results are presented in Table 5.2, Table 5.3 and Table 5.4, where it is evident that our network achieves state of the art performance with significantly fewer trainable parameters.

Model	Number of Trainable Parameters	Accuracy [14]
DeepConvNet [62]	226,677	**82.0% ± 0.40%**
ShallowConvNet [62]	107,922	79.7% ± 0.35%
EEGNet [63]	2,986	80.0% ± 1.00%
EEG-Inception [64]	22,386	80.5% ± 0.35%
BrainWaveScattering Net (our architecture)	6,050	**81.5% ± 0.13%**

Table 5.2: Performance of generic (trained across-subjects) models of MI-classification (Left / Right hand) task in PhysioNet for DeepConvNet, ShallowConvNet, EEGNet, EEG-Inception and BrainWave-Scattering Net (ours).

Model	Number of Trainable Parameters	Accuracy [14]
DeepConvNet [62]	180,402	63.1% ± 0.30%
ShallowConvNet [62]	193,442	61.5% ± 1.30%
EEGNet [63]	3,354	61.0% ± 1.00%
EEG-Inception [64]	26,742	62.2% ± 2.00%
BrainWaveScattering Net (our architecture)	6,300	**65.2% ± 0.70%**

Table 5.3: Performance of generic (trained across-subjects) models of MI-classification (Right Hand / Foot) task in BCI IVa dataset for DeepConvNet, ShallowConvNet, EEGNet, EEG-Inception and BrainWave-Scattering Net (ours).

Model	Number of Trainable Parameters	Accuracy [14]
DeepConvNet [62]	171,302	71.56% ± 0.24%
ShallowConvNet [62]	34,282	73.11% ± 0.4%
EEGNet [63]	1,962	72.05% ± 0.28%
EEG-Inception [64]	20,154	72.3% ± 1.06%
BrainWaveScattering Net (our architecture)	2,166	**73.14% ± 0.14%**

Table 5.4: Performance of generic (trained across-subjects) models of MI-classification (Left Hand / Rest) task in MyBrainCommands dataset for DeepConvNet, ShallowConvNet, EEGNet, EEG-Inception and BrainWave-Scattering Net (ours).

5.4.2 Personalised Models of Motor-Imagery Recognition

In this results section, we present our attempts to build personalised (trained within a subject) models of MI-recognition. This is actually the most popular strategy for overcoming the problem of inter-subject variability when building brain activity decoders (e.g. [179], [180] and [181]). However, this strategy is not applicable to deep learning models that require thousands of trials as training data, and labeled EEG-signals still remain scarce at the level of an individual. On the other hand, BrainWave-Scattering Net is an architecture that allows us to train personalised models without facing training instability and the problem of overfitting. We managed to craft personalised models not only in the case of BCI IVa dataset (where 280 trials were available

[14] ±% refers to the rounded standard deviation across multiple 10-fold cross-validation experiments

per subject), but even in the case of MyBrainCommands and PhysioNet data (where only 54 and 45 trials were available per subject respectively). Table 5.5 include the performance of all personalised MI-decoders, trained/validated based on a 10-fold cross-validation scheme. Table 5.6[4] and Table 5.7[4] include the performance of a randomly selected subset of personalised MI-decoders, trained/validated based on a 5-fold cross-validation scheme in the case of MyBrainCommands dataset and on leave-one-trial-out cross-validation scheme in the case of PhysioNet dataset, while the full lists are available in the Appendix.

Model	Subject aa	Subject al	Subject av	Subject aw	Subject ay
EEGNet [63]	61.1% ± 7.50%	70.3% ± 9.80%	55.3% ± 8.50%	57.9% ± 11.30%	58.9% ± 13.30%
BrainWave-Scattering Net (ours)	78.9% ± 6.20%	92.2% ± 4.20%	60.3% ± 9.20%	86.7% ± 4.30%	85.7% ± 3.10%

Table 5.5: Performance of personalised (subject-specific) models of MI-classification (Right Hand / Foot) task in BCI dataset IVa for EEGNet and BrainWave-Scattering Net (ours).

Model	Subject 7	Subject 11	Subject 14	Subject 94	Subject 115
BrainWave-Scattering Net (ours) [15]	82.9%	73.8%	92.7%	68.2%	77.8%

Table 5.6: Performance of personalised (subject-specific) models of MI-classification (Left Hand / Rest) task in MyBrainCommands dataset for BrainWave-Scattering Net (ours).

Model	Subject 1	Subject 7	Subject 29	Subject 33	Subject 40	Subject 85	Subject 101
BrainWave-Scattering Net (ours) [15]	68.8%	91.1%	93.3%	57.7%	73.3%	86.6%	62.2%

Table 5.7: Performance of personalised (subject-specific) models of MI-classification (Right / Left hand) task in PhysioNet dataset for BrainWave-Scattering Net (ours).

5.4.3 Interpretability

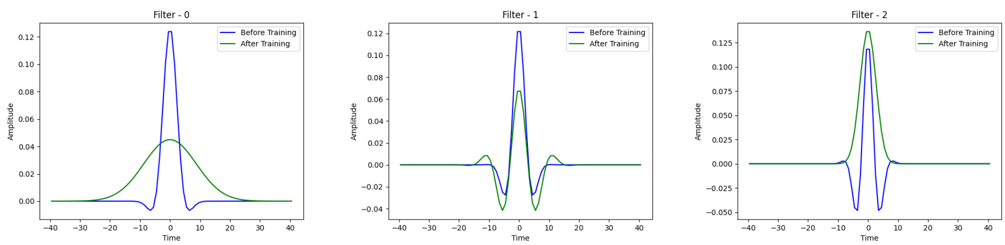

Figure 5.3: Gabor wavelets before and after training in the generic BrainWave-Scattering Net for the Left / Right Hand classification task in PhysioNet.

[15] Only the performance of BrainWave-Scatteting Net is included. The rest baseline networks could not be trained stably.

BrainWave-Scattering Net is a lightweight deep neural network compared to other well-established state-of-the-art architectures in the area of BCIs. Nevertheless, it achieves similar or even better performance in the tested MI-recognition tasks. In the previous section, it is shown that our network performs similarly, or better, to EEG-Inception and DeepConvNet in building generic MI-decoders in two distinct datasets, while having around $\frac{1}{4}$ of EEG-Inception's and $\frac{1}{30}$ of DeepConvNet's total trainable parameters (see Table 5.2).

Since the network is inspired by Joint Wavelet Scattering Transform [163], it inherits not only its stability but also its interpretability properties as well. More specifically, the first layer of the network consists of trainable Gabor wavelets. By allowing the frequency and bandwidth parameters of the Gabor wavelets to be trainable, the network gets adapted to variable EEG signals and extracts the most relevant information. This data-driven approach increases the explainability of the model since the first layer can now offer useful interpretable insights of the extracted features compared to the "black-box" approach used in other networks in the field. In Figure 5.3, different Gabor wavelets of the first layer of BrainWave-Scattering network are displayed before training (during initialization) as well as after training. These trainable Gabor wavelets are deployed in the first order wavelet scattering layer of the network (Block 1) and they are attested to be more informative compared to the corresponding first layer of other architectures (we refer the reader to Figures 5.6 for the learned temporal first layer filters of EEGNet [63] trained in the same motor-imagery tasks).

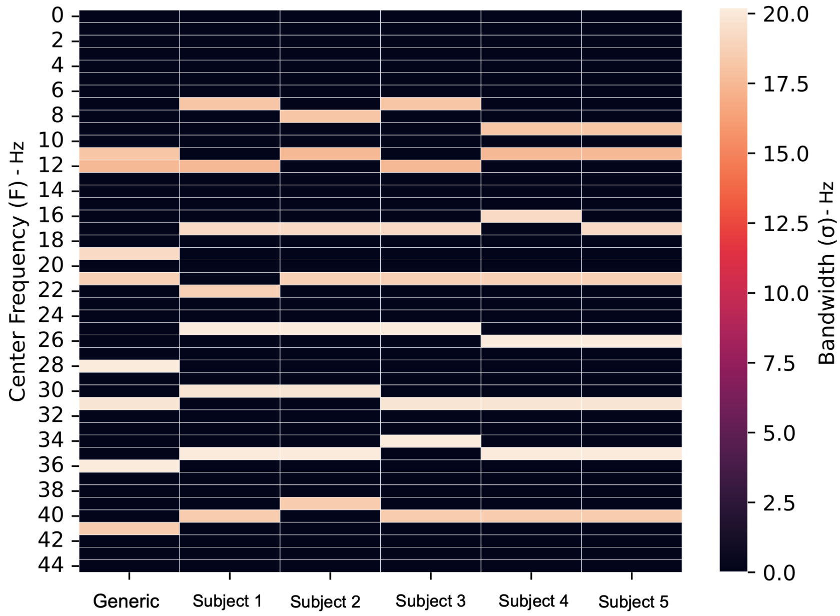

Figure 5.4: Gabor wavelets parameters for the generic and all the personalised models trained in the Right Hand/Foot classification task based on BCI Dataset IVa.

A second example of the interpretability offered by our network is provided via Figure 5.4, where the center frequencies and bandwidths of the trained first-layer filters are shown for the generic and the personalised models built based on the BCI dataset IVa. Here, it is shown how the Gabor filters have chosen different center frequencies in the generic model and the personalised models. In the latter case, the different frequencies is another evidence of the inter-subject variability. This can be an invaluable asset to neuroscientists who are familiar with the concept of brain-rhythms and accustomed to work with canonical frequency bands (or attempt to defined them with data-driven methodologies).

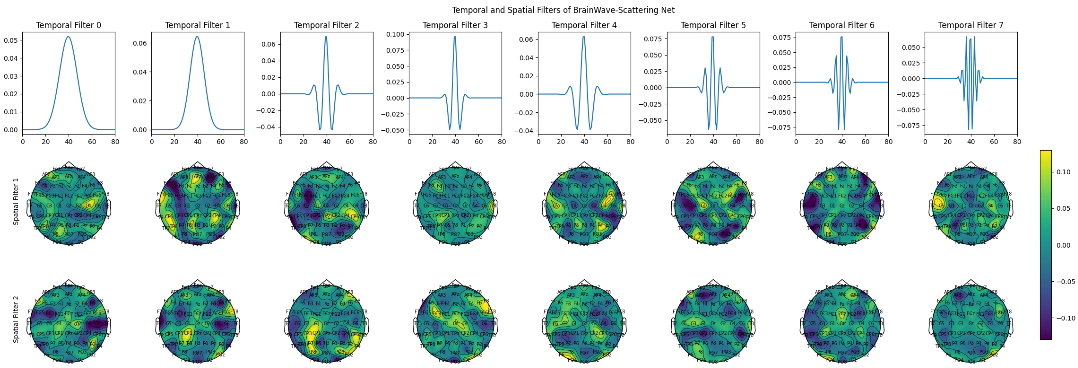

Figure 5.5: Learned temporal filters and their two associated learned spatial filters in the generic BrainWave-Scattering Net for the Left / Right Hand classification task in PhysioNet.

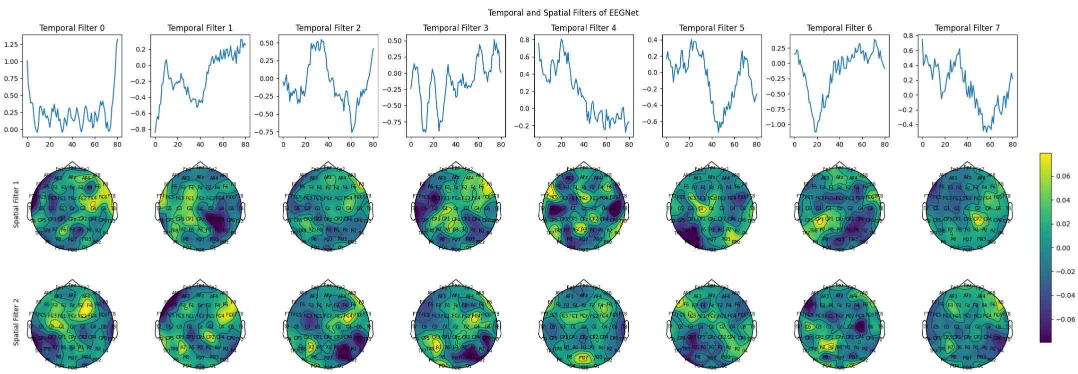

Figure 5.6: Learned temporal filters and their two associated learned spatial filters in the generic EEGNet [63] for the Left / Right Hand classification task in PhysioNet.

Finally, our architecture offers insights on the brain areas used with its trainable spatial filters in Block 4. An example of the spatial filter weights is provided in Figures 5.5 and 5.6.

5.4.4 Training time

Another great advantage of our architecture is the short training time it requires to achieve state-of-the-art performance in the tested MI-recognition tasks. Architectures like EEGNet [63] and EEG-Inception [64] have all their network's parameters trainable allowing a greater degree of freedom to the network. Therefore, as it shown in Figure 5.7, they are quickly adapted to the data achieving considerably high validation performance. In contrast, BrainWave-Scattering Net starts with almost random performance in the binary task. As described above, a learning scheduler is used for BrainWave-Scattering Net which consists of a large initial learning rate followed by a small rate at the end of the training phase. As it is displayed in Figure 5.7, the initial large learning rate assists the network's Gabor filters to quickly adapt to the data and the motor-imagery tasks while the small learning rate is primarily used for network's fine-tuning.

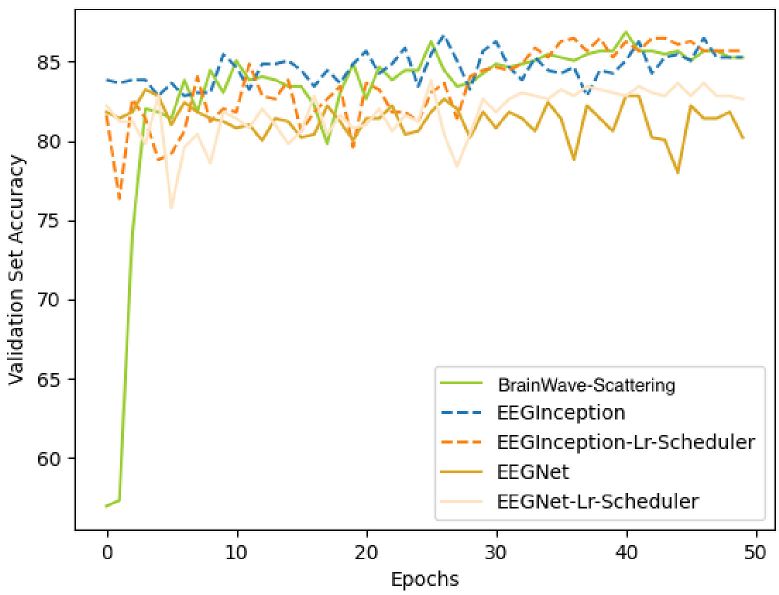

Figure 5.7: Per-fold training time for generic EEGNet, EEG-Inception and BrainWave-Scattering Net for the Left / Right Hand classification task in PhysioNet. All architectures are trained and tested in the same partition of the dataset.

Model	Number of Trainable Parameters	Training Time
EEGNet [63]	2,986	3:30 minutes
EEGNet [63] with learning rate scheduler	2,986	**3:27 minutes**
EEG-Inception [64]	22,386	8:04 minutes
EEG-Inception [64] with learning rate scheduler	22,386	7:56 minutes
BrainWaveScattering Net (our architecture)	6,050	**3:27 minutes**

Table 5.8: Per-fold training time for generic EEGNet (with and without learning rate scheduler), EEG-Inception (with and without learning rate scheduler) and BrainWave-Scattering Net (ours) for the Left / Right Hand classification task in PhysioNet.

Table 5.8 demonstrates that BrainWave-Scattering Net has similar performance to EEG-Inception [64] and it has the same training time as EEGNet [63], which it outperforms. For the training time tests, all architectures were tested in the same partition of the PhysioNet dataset and in the same NVIDIA TITAN Xp Graphics Processing Unit (GPU). Since BrainWave-Scattering Net uses a learning rate scheduler, for completeness in these results, EEGNet [63] and EEG-Inception [64] were trained in two setups: 1) with the same learning rate scheduler and 2) with a constant learning rate.

5.4.5 Transferability to lower density sensor arrays

The different way in which information from different sensors is combined during its flow through the BrainWave-Scattering Net allows us to use pre-trained models on dense setups into lower density sensor arrays while managing to maintain an accuracy close to the original performance.

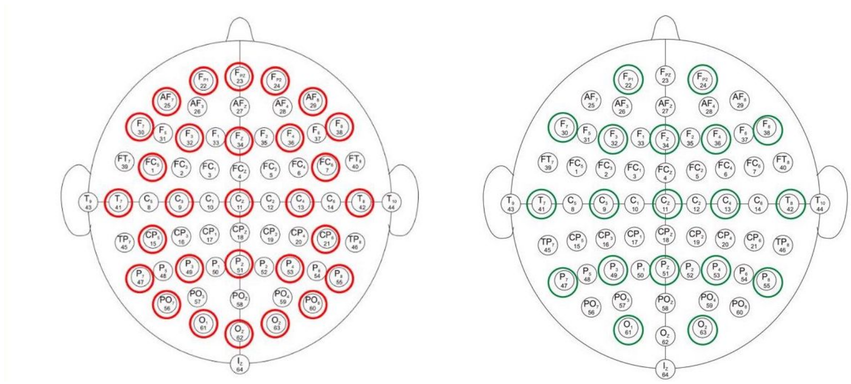

Figure 5.8: Schematic diagram of the original sensors used in the PhysioNet dataset and the selected lower dense sensor arrays: 29 sensors (in red) and 19 sensors (in green) - [182] [66].

For these experiments, the networks were trained first in the MI-recognition task of PhysioNet dataset following the cross-validation scheme described in section 5.4.1 using all available 64 EEG sensors / channels. When testing the trained generic model, only either 29 or 19 of the original input EEG signals (electrode configurations that correspond to EEG hardware devices on the market) distributed across the whole head were kept (see Figure 5.8) and the missing sensors were replaced with signals of zeros.

From the results displayed in the following table (Table 5.9), it is shown that BrainWave-Scattering Net model can be transferred more easily between different density sensor arrays without re-training.

Model	64-Channels	29-Channels	19-Channels
DeepConvNet [62]	**82.4%**	73.0%	69.0%
ShallowConvNet [62]	80.0%	67.0%	60.0%
EEG-Inception [64]	81.0%	69.0%	58.0%
BrainWave Scattering Net (our architecture)	81.5%	**76.0%**	**70.0%**

Table 5.9: Performance of generic models of MI-classification (Left / Right Hand) task in PhysioNet tested on a number of different sensors arrays.

5.4.6 Robustness to sensor removal

An additional element, that distinguish the BrainWave-Scattering Net architecture from other state-of-the-art networks in the field, is the way in which information from different sensors is combined during its flow through the network. In architectures like DeepConvNet [62], Shallow-ConvNet [62], EEGNet [63] and EEG-Inception [64] the combination of knowledge across the EEG channels takes place in the first layers of the network, while BrainWave-Scattering Net processes each channel separately and captures the spatial filters at the very end of the network. This fundamental difference provides an additional layer of robustness to our network during inference when some of the input EEG signals are tampered - for example several EEG sensors of a BCI headset are faulty.

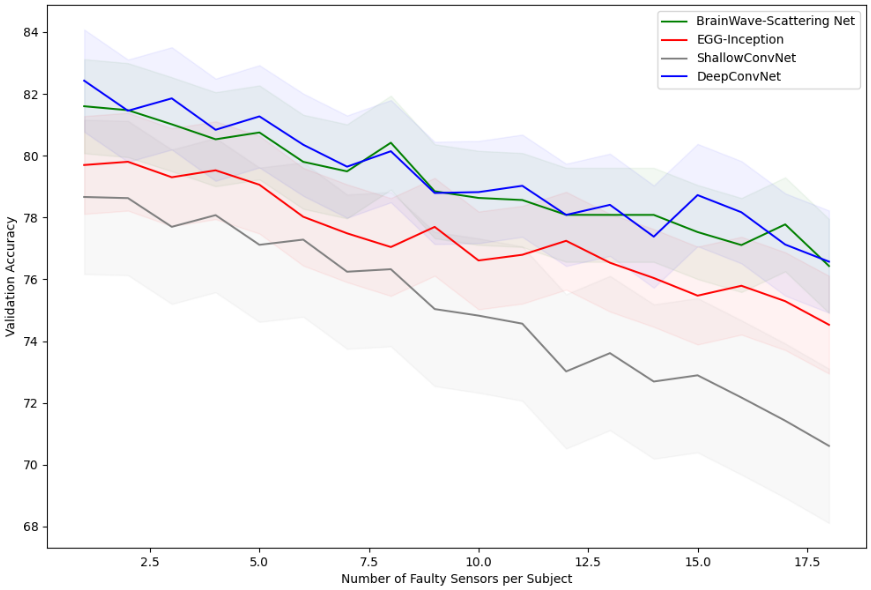

Figure 5.9: Performance of generic models of MI-classification (Left / Right Hand) task in PhysioNet (trained across-subjects and tested on subjects with a number of faulty sensors) for DeepConvNet (Blue), ShallowConvNet (Gray), EEG-Inception (Red) and BrainWave-Scattering Net (Green).

To simulate this "tampered device" scenario, the networks were first trained in the MI-recognition task of PhysioNet dataset following the cross-validation scheme described in section 5.4.1 using all available 64 EEG sensors / channels. When testing the trained generic model, a number of sensors was randomly chosen for each subject in the testing fold and their values were replaced with small random Gaussian noise. As it is evident in Figure 5.9, our architecture maintains an accuracy close to the original performance when a reasonable amount of sensors has failed.

5.4.7 Model exploration

One of the core advantages of our architecture is the introduction of a trainable element in the standard wavelet scattering pipeline. As mentioned previously, the scattering transform was first introduced in [169] to build a signal representation that is invariant to transformations (such as translations and diffeomorphisms) while preserving most of the signal's information. Although different architectures have already made use of various trainable deep wavelet networks ([165],[164]), our model is oriented around the field of Motor-Imagery and Brain-Computer Interfaces (BCIs), although the core idea can be applied in other domains of signal processing as well. In this section, we describe some experiments that led us to the final design of our architecture. By comparing and contrasting standard wavelet techniques with their equivalent data-driven approaches, we demonstrate the importance of the trainable element in our architecture.

Experiment 1 - Fully Trainable Scattering

In this experiment, the performance of a network with fixed wavelet filters is compared with the performance of a network with trainable filters. For the fixed network, Kymatio [183] is utilized which implements wavelet scattering transforms as convolutional networks with fixed wavelet filters. Kymatio takes a 1D signal as an input and outputs the second order wavelet scattering coefficients according to the following equation [183]:

$$Sx = ||x(t) * \psi_{\lambda_1}(t)| * \psi_{\lambda_2}(t)| * \phi(t) \tag{5.18}$$

where $*$ denote the convolution operator, $\psi_{\lambda_1}(t)$ and $\psi_{\lambda_2}(t)$ are Gabor wavelets and ϕ the average filter on the time domain. In this experiment, Kymatio was set to extract 7 second order wavelet coefficients and reduce the time analysis of the EEG signals by 6 in order to keep the number of network's paramaters at a low level. During the computation of the 7 second order wavelet coefficients, Kymatio produced 5 first order wavelet coefficients and downsampled intermediate results to reduce computational load as in [184]. Passing each EEG channel through Kymatio's network, a 3D tensor was generated with dimensions (EEG Channels $\times \frac{Time}{6} \times 7$). For the network with fixed wavelet filters, this extracted 3D Tensor was passed through the Block 4 of BrainWave-Scattering Net.

An equivalent fully-trainable network consists of a first layer with a fully-trainable convolutional kernel of size (1, $\frac{F_s}{2}$) where F_s is the sampling frequency followed by a modulus operator and average pooling of 4. Since the sampling rate of the signal is reduced by 4, the second layer consists of another fully-trainable convolutional kernel with size (1, $\frac{F_s}{8}$) followed by a modulus operator and average pooling of 2. To perform the time averaging operation, the network utilizes depthwise convolutions with stride of 2 and fully-trainable convolutional kernel with size (1, $\frac{F_s}{16}$) since the sampling rate of the signal is totally reduced by 8. These cascade of operations generated an equivalent trainable 3D tensor with dimensions (EEG Channels $\times \frac{Time}{6} \times 7$). Lastly, this trainable 3D Tensor was passed through the Block 4 of BrainWave-Scattering Net.

From the results displayed in Table 5.10, it is shown that the network with trainable layers outperforms the network with fixed Gabor filters showcasing that a fully-trainable data-driven scattering pipeline overpasses a standard wavelet scattering network.

Network	Trainable Second Order Scattering Net	Non-Trainable Second Order Scattering Net
PhysioNet [66]	**68.0%**	63.0%
BCI IVa [67]	**63.0%**	56.0%

Table 5.10: Performance of generic (trained across-subjects) models of MI-classification tasks (PhysioNet - Left / Right hand and BCI IVa - Right Hand / Foot) for a trainable second order scattering network with its non-trainable equivalent.

Experiment 2 - Trainable Gabor Filter

As mentioned previously, in order to increase the explainability properties of our architecture, the first layer of the network is restricted to follow the real part of the Gabor wavelet format. Generally, Gabor filters are parametrized by their center frequencies and bandwidths. In the experimental part, the center frequencies were initialized in the range of interest for motor-imagery tasks [8,40Hz] and the bandwidth was treated as a hyperparameter specifically fine-tuned for each dataset and task. And the question arises: is it necessary for these Gabor filters to be trainable even though they have been carefully initialized for a MI task ?

Network [16]	BrainWave-Scattering Net (Trainable)	BrainWave-Scattering Net (Non-Trainable)
PhysioNet [66]	**81.5% ± 0.13%**	79.0% ± 1.00%
BCI IVa [67]	**65.2% ± 0.70%**	61.0% ± 1.00%

Table 5.11: Performance of generic (trained across-subjects) models of MI-classification tasks (PhysioNet - Left / Right hand and BCI IVa - Right Hand / Foot) for BrainWave-Scattering with trainable and non-trainable Gabor filters in its first layer.

To answer this question in this experiment, two versions of the BrainWave-Scattering Net architecture were tested: one with trainable Gabor filters (trainable frequency and bandwidth) and one with fixed Gabor filters (fixed frequency and bandwidth). In both cases, the Gabor filters were initialized with the same values. From the results displayed in Table 5.11, it is shown that the network with trainable frequencies and bandwidths can lead to a greater performance than the network with carefully crafted Gabor filters, further backing the claim that a data-driven pipeline can outperform a standard wavelet scattering network.

Experiment 3 - Discriminability Index

Finally, the performance of the derived personalised BrainWave-Scattering Net models was compared with the potential to build reliable "classical" machine learning models for the same subjects, based on spectro-temporal features derived via standard Continuous Wavelet Transform (CWT). To this end, we worked independently for each subject and performed the following discriminative analysis procedure, which was adapted from ([180] and [181]) for a similar classification task (approached via a different single-trial representation scheme). Single-trial based scalograms were first derived (i.e. feature extraction step), based on analytic Morse wavelet, for every sensor and then compared statistically the distribution of the absolute values of the CWT at each point within the frequency-time domain. This step resulted to discriminability scores stored in a 3D tensor (Sensors $\times$ Frequency $\times$ Time), that could be used for identifying the most informative spectro-temporal features for the MI-classification task (i.e. feature selection step). A summary of this tensor was derived by first integrating over the frequency-time dimensions the sensor-dependent scores and then by identifying the overall maximum across sensors. In this way, a single discriminability index was associated with each subject, reflecting the difficulty to tackle the left/right hand classification task based on the CWT-representation of the recorded brainwaves. Figure 5.10 enables the direct contrast between the performance of our network and the derived discriminability index. It becomes evident that personalised performances follow mostly the same trend as the wavelet-based

[16] ±% refers to the rounded standard deviation across multiple 10-fold cross-validation experiments

discriminability index. Moreover the correlation between the two sets of measurements is strong and very significant (correlation coefficients: ρ=0.57, P-value=3.4259×10^{-10}). This clearly indicates that our architecture, although contains trainable elements, has its roots in wavelet analysis by following the wavelet scattering transform pipeline. In many cases, though, the personalised performance is considerably high compared to the corresponding discriminability index proving that the trainable adaptable element of the network can outperform standard wavelet analysis.

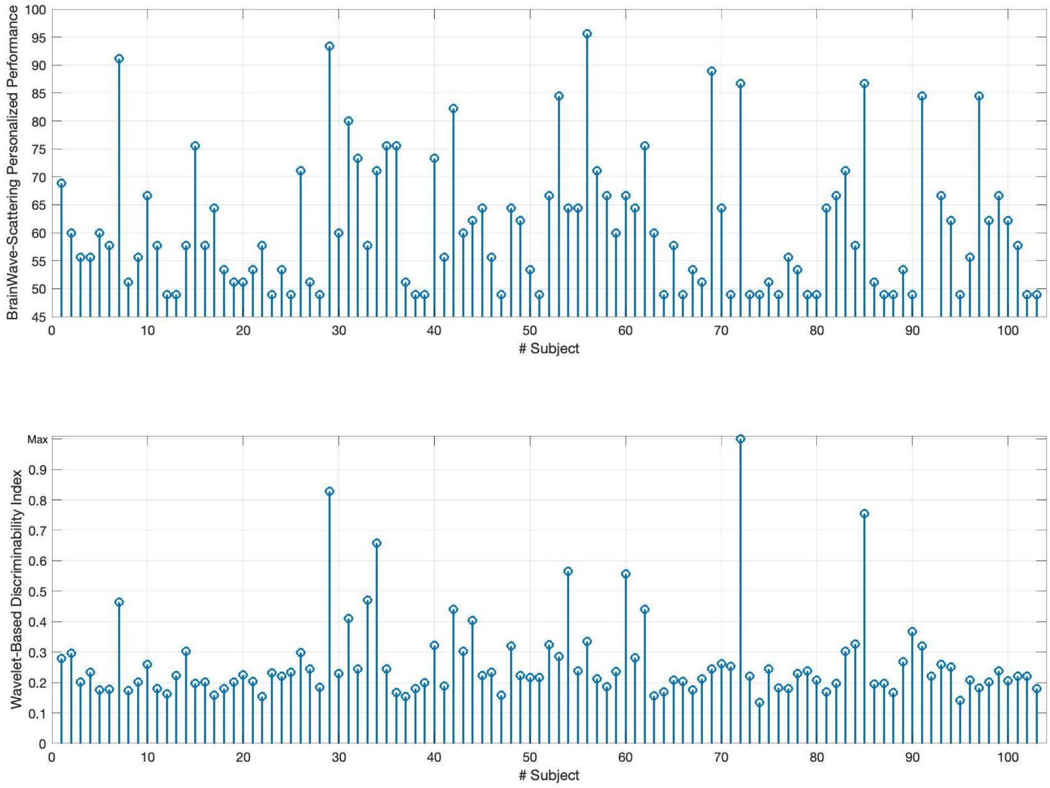

Figure 5.10: Subject-wise comparison between BrainWave-Scattering personalised performance and wavelet-based discriminability index. The horizontal axis represents the subjects while the vertical axis represents either the personalised performance (Top) or the wavelet-based discriminability index (Bottom).

5.4.8 Ablation Studies

Different Learning Rate Schedulers

As stated previously, since BrainWave-Scattering Net is lightweight, different training settings were used to ensure stability during training. A learning rate was used to assist the Gabor filters to quickly adapt to the training data. In this section, the effect of the same learning rate scheduler to other networks (both deep and shallower) was investigated. As it is shown in the following tables, the learning rate scheduler does not influence the other baseline lines by much while it can severely impact BrainWaveScattering-Net.

Network	Accuracy without lrs	Accuracy with lrs
DeepConvNet [62]	82.0%	82.5%
ShallowConvNet [62]	79.7%	80.0%
BrainWave-Scattering Net (our architecture)	79.1%	81.5%

Table 5.12: Per-fold training time for generic DeepConvNet (with and without learning rate scheduler), ShallowConvNet (with and without learning rate scheduler) and BrainWave-Scattering Net (with and without learning rate scheduler) for the Left / Right Hand classification task in PhysioNet. lrs stands for learning rate scheduler.

Network	Accuracy without lrs	Accuracy with lrs
DeepConvNet [62]	63.1%	63.9%
ShallowConvNet [62]	61.5%	61.0%
BrainWave-Scattering Net (our architecture)	62.1%	65.2%

Table 5.13: Per-fold training time for generic DeepConvNet (with and without learning rate scheduler), ShallowConvNet (with and without learning rate scheduler) and BrainWave-Scattering Net (with and without learning rate scheduler) for the Right Hand / Foot classification task in BCI IVa dataset. lrs stands for learning rate scheduler.

Different Bandwidth Values

In this section, we tried to determine the importance of the initialization values of the Gabor filters in our architecture. As stated previously, the initial values should be treated during the hyper-parameter tuning phase of the network. In addition, as described in [164], a reasonable range for the σ bandwidth values is $[4\sqrt{2log2}, 2W\sqrt{2log2}]$ "such that the full-width at half-maximum of the frequency response is within 1/W and 1/2", where W is the wavelet kernel window. Using the BCI IVa dataset [67], generic (trained across-subjects) BrainWaveScattering-Net models were trained by initializing the Gabor filters either with the same value (results are shown in Figure 5.11) or with randomly sampled values from the above mentioned range (the average of 10 different cross-validation performances when Gabor filters were initialized with randomly sampled bandwidth values is 63.9%). As it is evident from Figure 5.11, the network's performance can be influenced by this initialization choice. It is recommended that the Gabor filters should be initialized with a common value since that choice reduces the fine-tuning complexity and gives the same advantage to all Gabor filters to increase or decrease their respective bandwidths during training depending on the frequency range that they find important for the motor-imagery task they learn.

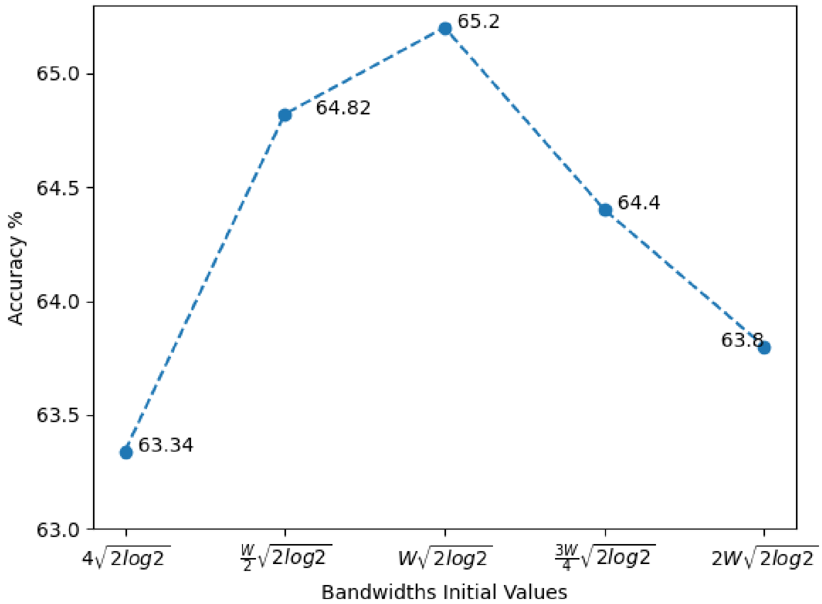

Figure 5.11: Performance of generic (trained across-subjects) models of MI-classification (Right Hand / Foot) task in BCI IVa dataset for BrainWave-Scattering Net. Blue points: Cross-validation performance when Gabor filters are initialized with the same bandwidth values.

5.5 Improving personalised Performance

In this section, a subject selection framework is introduced for augmenting the data available to craft a personalised model, by incorporating additional signals from subjects with similar brain activations. The overall framework utilizes brain signal analytics to compare responses at individual sensor levels (based on the wavelet transform) and employs a graph-based approach to analyze discriminative patterns across subjects, resulting in a semantic map facilitating the identification of relevant data for training personalised models. This selection framework demonstrates enhanced performance for personalised MI-BCIs, particularly benefiting subjects who initially exhibited low personalised performance. In the experiments, mainly BrainWave-Scattering Net is utilized, as a state-of-the-art lightweight deep learning architecture for personalised MI-BCI decoders and an architecture with roots to wavelet scattering transform.

5.5.1 Subject Selection Framework via Wavelets and Graph Diffusion

Enhancing the efficacy of personalised deep learning-based MI-BCI decoders through targeted subject selection during training necessitates the adoption of an appropriate metric to assess subject similarities. This selection framework leverages continuous wavelet transform (CWT) and graph diffusion distances.

Let $\mathbf{X}_i(c,t) \in \mathbb{R}^{C\times T}$ denote each single-trial EEG signal recorded with C channels that comes with the label $y_i \in \mathbb{R}$. For each subject in the dataset, and for each sensor independently, all the single-trial scalograms are first derived, using CWT based on analytic Morse wavelet. Then the distribution of moduli at each point in the time-frequency domain is statistically compared using two-sample unpaired Wilcoxon test [185]. The resulting discriminability scores $S(c,f,t) \in \mathbb{R}^{C\times F\times T}$ express the differences between the contrasted conditions (e.g. left vs right hand MI) for each sensor. Integration over the frequency-time dimensions yields a discriminability profile $\mathbf{p}(c) \in \mathbb{R}^{C}$ for each subject.

To systematically compare the discriminability profiles across subjects, the following graph-theoretic perspective is adopted so as to accommodate for the fact that these are patterns over a irregular domain (i.e. the sensor array). To this end, each profile is considered as a graph signal [186] and graph-diffusion distance is employed as a means to express their pairwise similarity relationships.

Using the Euclidean distance of spatial coordinates of the sensors, the adjacency matrix $\mathbf{A} \in \mathbb{R}^{C\times C}$ of the graph as well as its corresponding Laplacian $\mathbf{L} \in \mathbb{R}^{C\times C}$ are computed. Then the pseudo-inverse of the graph Laplacian $\mathbf{Q} = (\mathbf{I}+\mathbf{L})^{-1} \in \mathbb{R}^{C\times C}$ is obtained where $\mathbf{I}$ is the identity matrix. The graph diffusion distance between two subjects is computed using the equation:

$$d(s_1,s_2) = ||\mathbf{Q} \times (\mathbf{p}_{s_1} - \mathbf{p}_{s_2})||_2 = ||(\mathbf{I}+\mathbf{L})^{-1}(\mathbf{p}_{s_1} - \mathbf{p}_{s_2})||_2 \tag{5.19}$$

$(\mathbf{I}+\mathbf{L})^{-1}$ can be interpreted as the solution to a diffusion equation on the graph, formed by the EEG sensors, starting with an initial condition of $(\mathbf{p}_{s_1} - \mathbf{p}_{s_2})$. While the norm operation measures the magnitude or "spread" of this diffusion process. It provides a way to quantify how much the initial difference between $\mathbf{p}_{s_1}$ and $\mathbf{p}_{s_2}$ has diffused across the graph after applying the pseudo-inverse of the graph Laplacian. The rationale behind this computation is to counteract the spatial jitter in these profiles that reflect possible inter-subject anatomical variabilities. For a dataset with K unique subjects, applying our method yields a symmetric matrix $\mathbf{D} \in \mathbb{R}^{K \times K}$.

To represent the (dis)similarities conveyed in the symmetric matrix $\mathbf{D} \in \mathbb{R}^{K \times K}$ in an intelligible way, a classical multidimensional scaling (MDS) is performed. This results in a semantic map [187], in which each subject is mapped on a two-dimensional space in which the distances between points (subjects) approximate the (dis)similarities in matrix $\mathbf{D}$. For the subject selection step, the Euclidean distance is calculated between the subjects in this two-dimensional space and choose the N (where N acts as a hyper-parameter) neighbours to include in the training phase of the personalised deep learning-based MI-BCI decoder.

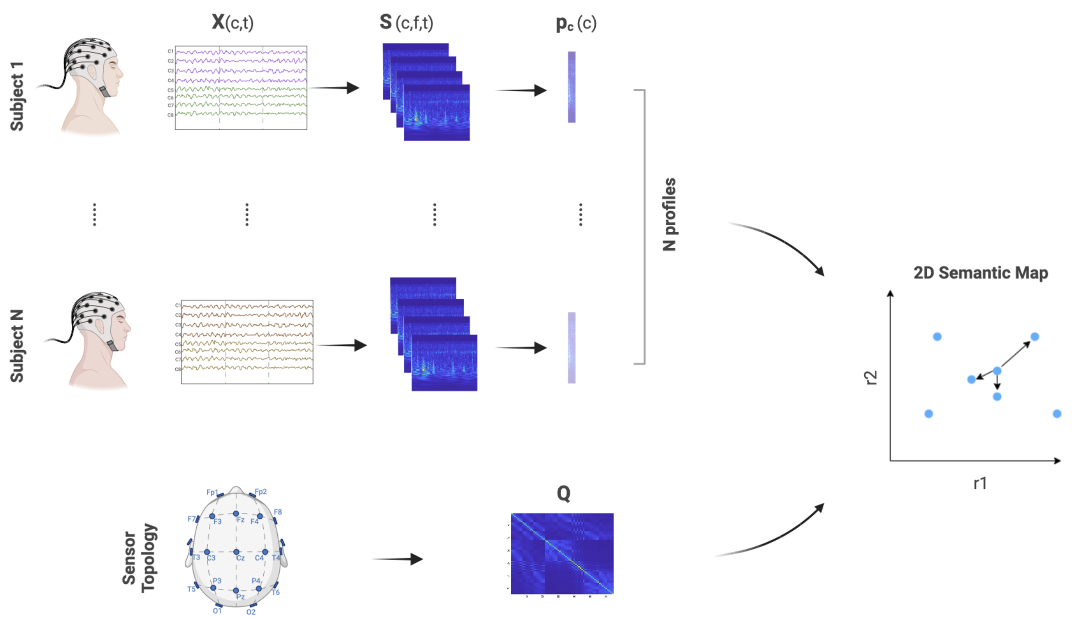

Figure 5.12: Our subject selection framework using wavelet transform and graph diffusion. For each subject, the discriminability scores S are calculated. Integration over the frequency-time dimensions yields discriminability profile $\mathbf{p}$ for each subject. From the sensor topology, the pseudo-inverse of the graph Laplacian $\mathbf{Q}$ is calculated. Using the $\mathbf{p}$ profiles and $\mathbf{Q}$, the symmetric matrix $\mathbf{D}$ is computed that contains the graph diffusion distances. Finally, the $\mathbf{D}$ is mapped onto a two-dimensional space and neighbouring subjects are selected.

5.5.2 Results

The subject selection framework was tested using BrainWave-Scattering Net, a state-of-the-art architecture for both personalised and generalized MI-BCI decoders, in a binary classification task (MI Left vs Right Hand), formed based on the publicly available MI dataset PhysioNet ([66]) as described in Chapter 2. Using our selection framework, the symmetric matrix $\mathbf{D} \in \mathbb{R}^{103\times 103}$ (excluding the 6 bad subjects) was computed and the two-dimensional discriminability map was generated, displayed in Figure 5.13. The personalised performance of the subjects on leave-one-trial-out cross-validation scheme was evaluated. We noticed that the subjects on the right-hand side of the discriminability map demonstrated higher personalised performance compared to the ones on the left-hand side.

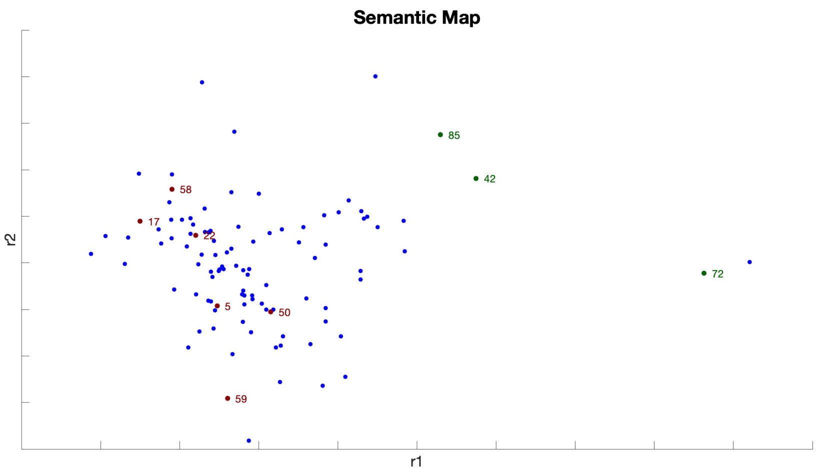

Figure 5.13: Two-dimensional semantic map for the 103 PhysioNet subject based on graph diffusion distances of our subject selection framework. Green subjects have personalised performance > 70% while red subjects have personalised performance < 70%.

The main objective with this subject selection framework is to investigate whether the personalised performance of bad performing subjects can be improved. To assess this hypothesis, 6 subjects exhibiting performance close to random levels during leave-one-trial-out cross-validation were identified (we refer the reader to the Appendix). For each of these subjects, we opted to incorporate 5 neighboring subjects ($N = 5$) in the training phase of their respective personalised models. To mitigate concerns regarding any potential performance enhancement being solely attributed to the increase in training data samples, the same personalised training was conducted using five randomly-selected non-neighboring subjects. Subsequently, the performance of these personalised models was compared with that of generic models trained on the remaining 102 subjects. As it is shown in Table 5.14, the models trained with our subject selection framework demonstrate increased personalised performance while in most cases outperforming their respective generic models.

Subject	Only Subject's Samples	With Neighbours	With Random	Generic
17	64.4%	**91.1%**	68.8%	86.6%
22	57.7%	**84.4%**	66.6%	73.3%
58	66.6%	**75.5%**	68.8%	64.4%
5	60.0%	**75.5%**	60.0%	**75.5%**
59	60.0%	**80.0%**	62.2%	73.3%
50	53.3%	**80.0%**	66.6%	**80.0%**

Table 5.14: Performance of personalised (trained and evaluated in a leave-one trial-out fashion) models of MI-classification (Left / Right hand) tasks in PhysioNet for BrainWave-Scattering Net using i) only subject's samples, ii) extra neighbours selected using our subject selection framework, iii) extra randomly selected subjects and iv) comparing with a generic model.

5.6 Conclusions

In this chapter, we introduce BrainWave-Scattering Net, a lightweight, fully-learnable architecture for EEG signal classification. Our network achieves state-of-the-art performance compared to other networks in the field while having only a fraction of their total trainable parameters and their training time. Since it is lightweight, it allows coherent training of personalised models. Moreover, it can demonstrate better interpretability properties. To the best of our knowledge, this is the first network that implements a trainable joint time-frequency scattering transform and the first of its kind to be used in the field of BCIs.

In addition, in this chapter, we also introduced a subject selection framework, harnessing wavelet transform and graph diffusion distances. Through experimentation, we showcased that this selection framework achieves increased personalised performances of subjects who initially exhibited poor performance. This advancement holds particular significance in medical applications, where personalised models play a pivotal role in achieving tailored and effective interventions. Potential next experiments will include evaluation on more MI datasets and MI tasks, training with other deep learning-based MI decoders (that allow training of personalised models with limited training data) as well as an ablation study of the influence of neighboring-selected subjects (N) in the results.

6
Improving Generalization of CNN-Based MI Decoders

Contents

In this chapter, we introduce a framework that tackles the issue of inter-subject variability in CNN-based BCI models. Initial results of this framework were published in **K. Barmpas**, Y. Panagakis, D. Adamos, N. Laskaris, and S. Zafeiriou, *"A causal viewpoint on motor-imagery brainwave decoding"*, in ICLR2022 Workshop on the Elements of Reasoning: Objects, Structure and Causality (2022) while an extended analysis of the framework's abilities was published in **K. Barmpas**, Y. Panagakis, S. Bakas, D. Adamos, N. Laskaris, and S. Zafeiriou, *"Improving generalization of CNN-based motor-imagery EEG decoders via dynamic convolutions"*, in IEEE Transactions on Neural Systems and Rehabilitation Engineering (2023).

6.1 Introduction

Guided by the causal framework of Chapter 4 and using the causal factorization for the MI classification problem (Figure 4.4), in this chapter we will focus mainly on the problems related to the anatomical differences of subjects $P(\mathsf{Z}|\mathsf{Y})$.

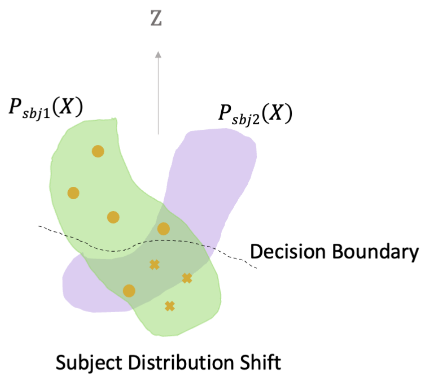

Figure 6.1: Key challenges in Machine Learning for a MI EEG classification task regarding the anatomical differences of subjects $P(\mathsf{Z}|\mathsf{Y})$. X represents input EEG signals and Z the true unobserved brain activity. $\bullet$ and $\times$ represent EEG signals of different labels. - Part of Figure 4.4.

Challenge: Each subject has a unique brain anatomy and functionality that results in polymorphous neural activity patterns when appeared in the surface observed EEG signal (e.g. [161], [162]).

Although many state-of-the-art deep learning architectures achieve impressive performance in MI classification tasks (e.g. [62], [63], [64]), they usually fail to tackle the problem of inter-subject variability [160], preventing the successful deployment of a previously trained MI classifier to new unseen subjects. Inter-subject variability is defined as the change in data distributions across different subjects: each individual has a unique brain anatomy and functionality that makes the discovery and exploitation of shared invariant features extremely difficult. In fact, these differences are so distinct that previous works have shown that the identification of a specific subject out-of-many is actually feasible (e.g. [188], [189], [190]). Therefore, modern DL-based BCIs tend to fail to generalize well in unseen subjects due to this type of data distribution shift. For many years, normalization techniques (e.g. [191], [192]) - data scaling using a mean and standard deviation - in conjunction with classical machine learning techniques have been considered the gold standard to solve the problem of inter-subject variability. With the advent of Deep Learning, methods like transfer learning have emerged in an effort to provide a solution (e.g. [193], [194], [195], [196]). In most of these methods, a small calibration set from the unseen subject is utilized to fine-tune parts of the pre-trained deep network architecture. In [196] only the last fully-connected layers are fine-tuned while the previous layers are frozen. In [193] some identified layers are fine-tuned to maximize knowledge transfer for MI classification. Although transfer learning has been proven to perform well, it still requires a calibration session in order to generalize well to unseen subjects. In the direction of zero-calibration networks, [142] proposes an adversarial inference framework that learns subject invariant features. In this chapter, we aspire to provide an alternative solution to the problem of inter-subject variability and enhance the above mentioned BCI deep architectures dynamically without the need of a calibration session.

In summary, this chapter's contributions are as follows:

1. A subject attention network based on learnable Gabor wavelets is introduced, which can accurately identify the different available subjects.

2. A framework based on dynamic convolutions is introduced, inspired by [197], which utilizes our subject attention network and with zero calibration is provably utilized to tackle the issue of inter-subject variability in the task of MI brainwave decoding according to our causal breakdown. More specifically, our causal analysis, described in Chapter 4, allows an evaluation setup to be designed which keeps all the identified distribution shifts intact but the inter-subject variability. Therefore, unlike other works in the area which claim improved cross-subject performance and often utilize a mixture of techniques like data augmentation (which can affect also other causal variables of interest), our framework is theoretically proven to target the problem of inter-subject variability through this specifically crafted evaluation setup.

The remainder of the chapter is organized as follows: Section 6.2 outlines our dynamic framework based on dynamic convolutions that improves the generalization of MI-BCI systems. Section 6.3 consists of the experimental part, where performance results and comparisons are detailedly presented. It also contains a discussion part which demonstrates the advantages and disadvantages of the dynamic framework. The last section 6.4 summarizes and concludes the work presented in this chapter.

The main objective in this chapter is to:

present a framework that improves the **generalization** of CNN-based MI decoders

6.2 Dynamic BCI Framework

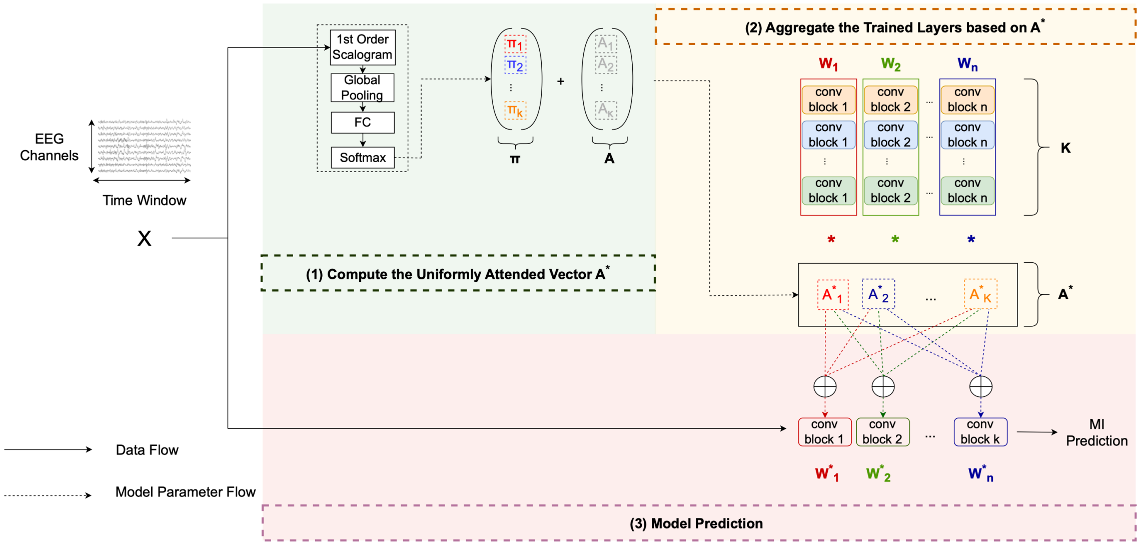

Figure 6.2: Dynamic convolution framework for BCI architectures. X represents input EEG signal trial. The K different subjects in the training set are represented by different colors in the convolutional blocks. Colored rectangles and arrows (namely green, red and dark blue) demonstrate the different blocks that are taken into account when computing the final convolutional blocks for the MI classification task.

In this chapter, we mainly focus on the challenge of subject distribution shift (or inter-subject variability). Using the causal breakdown described in Chapter 4, two publicly available MI datasets are used - which contain a large number of different subjects, are class balanced, have relatively enough trials per subject and all trials come from a single EEG recorder (within each dataset) - essentially solving all the identified challenges of Figure 4.4 but the subject distribution shift. In terms of the causal factorization (Eq. 4.7), the problem of inter-subject variability can be seen as a distribution shift $\mathcal{S}$ where:

$$P(\mathsf{X},\mathsf{Y},\mathsf{Z}) = P(\mathsf{X}|\mathsf{Z})P_{\mathcal{S}}(\mathsf{Z}|\mathsf{Y})P(\mathsf{Y}) \tag{6.1}$$

The dynamic framework can be applied to any established CNN-based MI-BCI architecture, resulting in performance increase. Inspired by [197], dynamic convolutions are utilized in the domain of MI brainwave decoding. Instead of having a BCI architecture that tries to discover a common latent space for all K subjects in the training set, K parallel trainable convolutional layers (corresponding to the K available training subjects) are used for each convolutional block of a CNN-based BCI network. Using a subject attention network that learns to distinguish between the available individuals, the subjects are separated from one another and essentially K parallel personalised models of the same BCI architecture are trained simultaneously, as illustrated in Figure 6.2.

The dynamic framework is inspired by the work of [197] in the field of Computer Vision, but it includes various modifications to address challenges apparent in the EEG domain. Although the complete framework will be detailedly described in the following subsections 6.2.1 and 6.2.2, these differences can be summarized as follows:

- Instead of fully trainable attention mechanisms, it utilizes our novel subject attention network (described in 6.2.1) which is inspired by Block 1 of BrainWave-Scattering Net (Chapter 5). It uses only trainable Gabor filters making it more lightweight and explainable than a shallow fully trainable neural network and it achieves very high performance in the subject identification task.
- Unlike [197] where there is an attention mechanism for each convolutional layer and these mechanisms are trained in an unsupervised manner, this framework uses only one attention mechanism for all convolutional layers, and with supervised training, it learns to distinguish between the available different subjects.
- The K number of parallel layers in the dynamic framework is not a tunable hyperparameter (like in [197]) but coincides with the number of available subjects in the training set.
- Instead of using the output vector of the attention mechanism as [197], this framework utilizes the "uniformly attended" vector **A*** (described in 6.2.2) in order to be more robust to the low Signal-to-Noise Ratio (SNR) of the EEG signal.

6.2.1 Attention Network

The first layer of our subject attention network is the first order wavelet scalogram of the input EEG signal. Mathematically, let $\mathbf{x}(t) \in \mathbb{R}^T$ denote a one-dimensional input EEG signal, where T is the number of EEG time points, and $\boldsymbol{\psi}_\lambda(t)$ be a wavelet. The 1st order scalogram is defined as $\mathbf{X}(\lambda, t) = |\mathbf{x}(t) * \boldsymbol{\psi}_\lambda(t)|$, where $*$ stands for the convolution operator. To perform this operation, the raw input signal from each EEG channel is convolved with a wavelet kernel with size $(1, W) = (1, \frac{F_s}{2})$ where F_s is the sampling frequency. This wavelet kernel follows the real Gabor wavelet format:

$$\psi_\lambda(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} \cos(2\pi\lambda t) \tag{6.2}$$

with t $\in [-\frac{W}{2}, \frac{W}{2}]$ and $\frac{1}{\sigma}$ denotes the bandwidth and λ the normalized frequency of the Gabor wavelet and these two properties are the only trainable parameters of this layer. During training, λ is restricted ($\lambda \in [0, \frac{1}{2}]$) to satisfy the Nyquist theorem. The three-dimensional (3D) tensor $X(c, \lambda, t) \in \mathbb{R}^{C \times F \times T}$ (where F is the number of Gabor filters and C the number of EEG channels) containing the first order wavelet scalograms $\mathbf{X}(\lambda, t)$ for each EEG channel is followed by a global average pooling across time and frequency. Finally, the resulted vector is passed through a fully-connected layer to compute the subject id vector $\boldsymbol{\pi}$.

6.2.2 Subject-Attended Dynamic Convolutions

The dynamic framework takes the EEG signal X as input and tries to learn both the correct MI task Y (estimate the conditional probability $P(Y|X)$) as well as the correct subject id π (estimate the conditional probability $P(\pi|X)$). The subject attention network and the K parallel convolutional layers are trained simultaneously using the following loss function:

$$Loss = (1 - acc) \times \ell_{Attention} + acc \times \ell_{MI} \tag{6.3}$$

where *acc* is the training accuracy of the subject attention network and ℓ denotes the cross-entropy function ($\ell_{Attention}$ for the subject attention network and ℓ_{MI} for the MI classification task). This loss function effectively enforces first the training of the subject attention network and, as the attention's accuracy increases, it switches its focus to train the parallel convolutional layers for the different MI tasks. As also suggested in [197], since softmax does not work well due to its near one-hot output, a large temperature in the softmax of the attention network is used during training in order to flatten the framework's attention, allow a broader gradient backpropagation and effectively assist in the subject attention network's training in the early epochs.

During inference, when an input EEG signal from a new unseen subject S_x is processed, it passes firstly through the attention network and the subject attention vector $\boldsymbol{\pi}$ is computed where $\sum_i \pi_i = 1$. We empirically observed that this vector is quite sparse, and if it was used during inference, only a handful of parallel convolutional layers would be utilized during the mixing. Instead, we would ideally like to use knowledge from all K individuals and "shift" the attention more to the most relevant subjects. To accomplish that, what we call the "uniformly attended vector" **A*** is computed. If there was no attention network, the K parallel convolutional layers would be mixed with a uniform factor $A_i = \frac{1}{K}$. To compute the "uniformly attended vector", the uniform attention vector **A** is combined with the subject attention vector $\boldsymbol{\pi}$ and the result is passed through a softmax activation to flatten the attention across all subjects - while maintaining the focus on the most relevant ones (we refer the reader to section 6.3.7 for a performance comparison between using $\boldsymbol{\pi}$ and **A*** as attention). Mathematically, this operation can be described as:

$$\underbrace{\begin{pmatrix} A_1 \\ A_2 \\ \ldots \\ A_k \end{pmatrix}}_{\mathbf{A}} + \underbrace{\begin{pmatrix} \pi_1 \\ \pi_2 \\ \ldots \\ \pi_k \end{pmatrix}}_{\pi} \xrightarrow{\sigma} \underbrace{\begin{pmatrix} A_1^* \\ A_2^* \\ \ldots \\ A_k^* \end{pmatrix}}_{\mathbf{A*}} \tag{6.4}$$

where σ denotes the softmax operation, **A** the uniform attention vector with $A_i = \frac{1}{K}$ and **A*** the "uniformly attended" vector where $\sum_i A_i^* = 1$. Let us denote with $\mathbf{W}_i^j$ the learned convolutional kernel of the network's i^{th} convolutional layer from the j^{th} parallel network and with $\mathbf{W}_i^*$ the dynamic convolutional kernel of the network's i^{th} layer, as illustrated in Figure 6.2. In our dynamic framework, we compute the dynamic convolutional as follows:

$$\mathbf{W}_i^*(x) = \sum_{k=1}^{K} A_k^*(x)\mathbf{W}_i^k \tag{6.5}$$

In other words, using the causal factorization (6.1), our dynamic framework estimate the probability $P_{S_x}(\mathsf{Z}|\mathsf{Y})$ of a new unseen subject S_x as the linear combination of K learned conditional probabilities. More specifically:

$$P_{S_x}(\mathsf{Z}|\mathsf{Y}) = A_1^* \times P_{S_1}(\mathsf{Z}|\mathsf{Y}) + A_2^* \times P_{S_2}(\mathsf{Z}|\mathsf{Y}) + \ldots + A_k^* \times P_{S_k}(\mathsf{Z}|\mathsf{Y}) \tag{6.6}$$

6.3 Experiments

To validate this dynamic framework based on our causal breakdown in Chapter 4, two publicly available MI datasets are used, namely **PhysioNet** [66] and **OpenBMI - MI** [68], as they were described in Chapter 2.

6.3.1 Subject Verification

The subject attention mechanism is a vital part in the dynamic framework. Therefore, its performance was evaluated separately first in order to ensure its ability to distinguish between the various available subjects in the two datasets. 10-fold trial-wise cross-validation was performed to measure its performance. Adam optimizer was used with learning rate of 0.01 for the first 30 epochs (to allow the Gabor filters to quickly adapt to the data) and 0.001 for the remainder 20 epochs. As shown in Table 6.1, the subject attention network performs sufficiently well in both datasets which makes it an ideal candidate for the attention mechanism in the dynamic framework.

Dataset	CV Average Accuracy[17]
PhysioNet (103 Subjects)	98.5% ± 0.13%
OpenBMI - MI (54 Subjects)	90.3% ± 0.07%

Table 6.1: Performance of Subject Attention Network (trained and evaluated using 10-fold cross-validation) in predicting the subject id in PhysioNet and OpenBMI - MI datasets. CV stands for Cross-Validation.

6.3.2 Comparison Between Standard and Dynamic Models

The dynamic framework was tested in four well-established BCI architectures, namely DeepConvNet [62], ShallowConvNet [62], EEGNet [63] and EEG-Inception [64] in the following MI tasks: for the publicly available MI dataset PhysioNet [66] one binary classification task (MI Left vs Right Hand) and a 3-class classification problem (MI Left Hand / Right Hand / Feet)) and for OpenBMI - MI [68] one MI binary classification task (MI Left vs Right Hand).

Model Setup	**Vanilla**	**Dynamic**
Optimizer	Adam	Adam
Training epochs	30	30
Learning Rate	0.001	0.01 (first 20 epochs) and 0.001 (last 10 epochs)
Number of Gabor Filter	-	8
Initialized Center Frequencies (Hz) of Gabor Filters	-	$\{8, 12, 16, 20, 24, 28, 32, 36\}$
Temperature used in the softmax	-	30

Table 6.2: Hyper-parameter choices for the experimental comparisons.

[17] ±% refers to the rounded standard deviation across 10 runs of 10-fold cross-validation experiments

As shown in Table 6.2, the standard networks were trained for 30 epochs with learning rate of 0.001 while their dynamic versions for 30 epochs - in the first 20 epochs with learning rate of 0.01, to assist the attention's Gabor filters to quickly adapt to the data, and 10 epochs with learning rate of 0.001 and frozen attention, to fine-tune to the MI task. In all cases, an Adam optimizer was used. Finally, a temperature of 30 was used during training in the attention mechanism as described in the previous section. The performances of the standard networks and their equivalent dynamic networks were evaluated in a leave-one-subject-out fashion (cross-subject performance) Table 6.3. Additionally, the p-value of paired t-tests between performance of standard networks and their equivalent dynamic networks was computed.

Dataset **Task**	PhysioNet MI Left / Right	PhysioNet MI Left / Right Hand / FeetHand	OpenBMI - MI MI Left / Right Hand
ShallowConvNet Dynamic ShallowConvNet	80.6 ± 11.4% **83.3 ± 12.7%** (102 / ≈0.0055 / 97.5)	66.3 ± 16.0% **69.0 ± 16.3%** (102 / ≈0.0043 / 95.4)	66.3 ± 11.1% **70.3 ± 11.1%**[18] (53 / 50.68)
DeepConvNet Dynamic DeepConvNet	82.5 ± 11.5% **83.5 ± 13.0%** (102 / ≈0.13 / 100.8)	67.1 ± 14.6% **71.3 ± 16.1%** (102 / ≈0.0000328 / 100.36)	71.7 ± 12.0% **73.1 ± 11.6%**[18] (53 / 52.47)
EEGNet Dynamic EEGNet	79.5 ± 12.4% **80.2 ± 13.5%** (102 / ≈0.26 / 79.8)	66.5 ± 13.2% **67.5 ± 15.4%** (102 / ≈0.15 / 72.2)	**74.9 ± 11.0%** 71.9 ± 12.1%[18] (53 / 41.42)
EEG-Inception Dynamic EEG-Inception	81.6 ± 11.8% **83.9 ± 11.9%** (102 / ≈0.0068 / 94.5)	67.6 ± 15.1% **71.4 ± 15.0%** (102 / ≈0.00025 / 92.84)	76.5 ± 10.8% **77.4 ± 10.0%**[18] (53 / 49.20)

Table 6.3: Performance of generic (trained and evaluated in a leave-one-subject-out fashion) models for DeepConvNet, ShallowConvNet, EEGNet, EEG-Inception and their Dynamic equivalent networks (ours). The K parameter used in dynamic models is coloured with violet. The p-value of paired t-tests between performance of standard and dynamic is coloured with gray. The ratio of trainable parameters ($\frac{Dynamic}{Standard}$) is coloured with blue.

6.3.3 Comparison With State-of-the-Art Approaches

Here, we are not only interested in comparing the models trained with the dynamic framework versus regularly trained CNN-based BCI architectures but also to compare the framework with other transfer learning approaches in the EEG domain. Therefore, the performances of the standard networks and their equivalent dynamic networks were evaluated in a leave-M-subjects-out fashion Table 6.4. Furthermore, the framework was compared with two other commonly used transfer learning EEG techniques: 1) an adversarial approach, namely [142], that (similarly to our approach) does not use a calibration set and 2) Euclidean alignment [141] that projects data into a domain-

[18] early stopping has been applied to some folds during the fine-tuning phase.

invariant space but it uses all the trials of a subject. The Euclidean alignment networks were trained similarly to their vanilla equivalent after performing the data projection for each subject. And the equivalent adversarial networks were trained with early stopping and adversarial regularization weight $\lambda = 0.005$ (hyperparameters taken from the original paper [142]). As it can be seen from Table 6.4, our method outperforms adversarial networks (a similar zero-calibration method) while it achieves the same or higher performance when compared with Euclidean alignment. It is worth mentioning though that Euclidean alignment uses all the trials of an unseen subject while the dynamic framework is dynamically adapted for each trial during inference.

Model **Dataset**	5-Fold CV ($K \approx 83$) PhysioNet	10-Fold CV ($K \approx 93$) PhysioNet	20-Fold CV ($K \approx 98$) PhysioNet	5-Fold CV ($K \approx 43$) OpenBMI - MI	Trials of Unseen Subject Used for Adaptation
Vanilla DeepConvNet	81.4 ± 1.3%	81.7 ± 2.9%	82.4 ± 5.5%	72.1 ± 2.3%	-
DeepConvNet with Euclidean Alignment	82.7 ± 1.2%	83.03 ± 3.2%	**83.2 ± 5.0%**	70.6 ± 2.0%	All Trials
Adversarial DeepConvNet [18]	81.4 ± 1.35%	81.8 ± 3.7%	81.9 ± 4.6%	66.6 ± 2.0%	-
Dynamic DeepConvNet (Ours)	**82.8 ± 2.03%**	**83.14 ± 3.9%**	83.2 ± 5.3%	**72.3 ± 2.3%**	1
Vanilla ShallowConvNet	79.7 ± 1.75%	80.4 ± 3.3%	80.95 ± 4.6%	62.1± 1.3%	-
ShallowConvNet with Euclidean Alignment	79.8 ± 2.0%	80.3 ± 3.5%	81.1 ± 4.3%	63.8 ± 3.8%	All Trials
Adversarial ShallowConvNet [18]	81.0 ± 1.26%	81.3 ± 3.7%	82.35 ± 4.65%	55.1± 0.9%	-
Dynamic ShallowConvNet (Ours)	**81.0 ± 1.06%**	**82.32 ± 2.86%**	**83.10 ± 4.7%**	**68.4 ± 2.9%**	1

Table 6.4: Performance of generic (trained and evaluated in a leave-M-subjects-out fashion) models of MI-classification (Left / Right hand) tasks in PhysioNet and OpenBMI-MI. CV stands for Cross-Validation across subjects

6.3.4 Calibration Methods

The performance of the calibrated networks (using a small calibration set of the unseen subjects to fine-tune the final classification layer) was evaluated as well. For a fair comparison, the last layer of the equivalent dynamic networks using the same calibration sets was also fine-tuned. As it is shown in Table 6.5, the calibrated dynamic models also outperform their equivalent vanilla calibrated networks.

Model	5-Fold CV ($K \approx 83$)	10-Fold CV ($K \approx 93$)	20-Fold CV ($K \approx 98$)
Calibrated DeepConvNet [19]	81.94 ± 1.95%	82.3 ± 3.0%	82.5 ± 5.26%
Calibrated Dynamic DeepConvNet [19]	**83.2 ± 1.5%**	**83.35 ± 3.35%**	**83.43 ± 4.8%**
Calibrated ShallowConvNet[19]	80.6 ± 1.2%	81.5 ± 3.5%	81.45± 4.6%
Calibrated Dynamic ShallowConvNet [19]	**82.65 ± 1.55%**	**83.9 ± 3.85 %**	**84.00 ± 5.2%**

Table 6.5: Performance of generic (trained and evaluated in a leave-M-subjects-out fashion) models of MI-classification (Left / Right hand) tasks in PhysioNet for Calibrated DeepConvNet and ShallowConvNet and their Calibrated Dynamic equivalent networks. CV stands for Cross-Validation across subjects.

[19] Average result of 10 calibrated models. Each model uses 20% of each testing subject's samples for calibration (chosen randomly) and 80% for testing.

6.3.5 Investigation of Negative Transfer Learning

Although the dynamic framework showed increased cross-subject performance as experimentally shown above, we wanted to investigate if there are any signs of negative transfer learning during the process. As it is shown in Figure 6.3, although there are limited cases of negative transfer learning, the vast majority is either marginally or significantly better compared to the vanilla architectures.

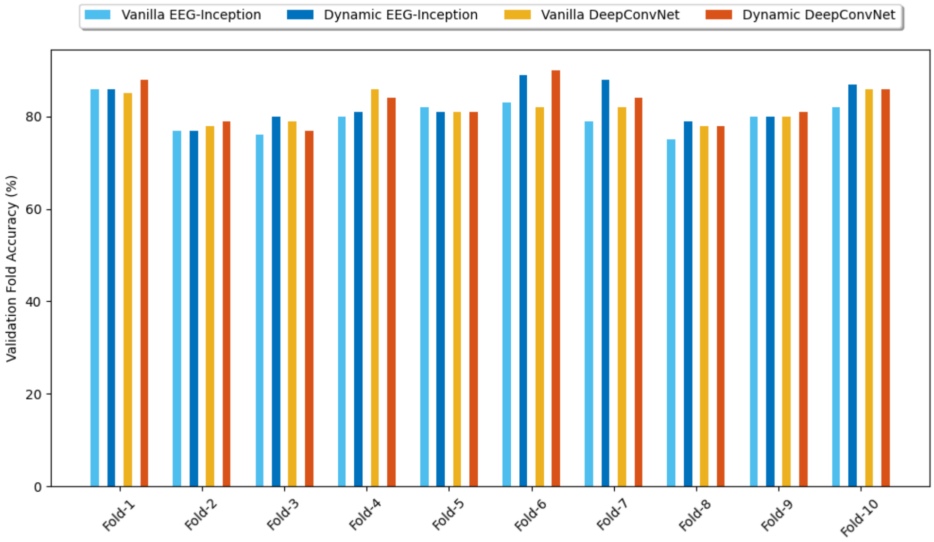

Figure 6.3: Per-fold comparison of the performance of vanilla architectures versus their equivalent dynamic networks (ours).

6.3.6 Limitations

The dynamic framework can be used in various CNN-based MI-BCI architectures to increase the cross-subject performance and can take the BCI community one step closer in tackling the problem of inter-subject variability as the experimental evaluations in the previous subsections illustrate. We expect this framework, with certain modifications, to be able to generalize well and get adapted to various BCI paradigms, not only MI. Investigating different BCI paradigms is beyond the scope of this Thesis where the causal analysis of the MI task is a core factor in ensuring that the dynamic framework tackles the targeted problem and there are no misleading performance increases. Extending the framework to different paradigms would require also a causal breakdown for these tasks using the causal framework of Chapter 4.

One limitation of this dynamic framework is the unavoidable increase in the number of trainable parameters (about $\times$ K where K is the number of available subjects in the dataset). Although the subject attention mechanism seems to identify well a large number of subject (e.g. 103 on PhysioNet), this increase in the number of trainable parameters might be a limiting factor in some cases especially if these models are deployed on real-life applications where devices have limited amount of memory storage. Fortunately, this tremendous increase in number of parameters does not translate to execution time. As it is shown in Table 6.6, there is a less than $\times$ K increase in terms of inference time cost.

Subject	Standard (ms)	Dynamic (ms)
Subject 1	561.9779	1146.8094
Subject 2	543.0056	1099.2773
Subject 3	563.3565	1139.8957
Subject 4	545.2918	898.6636
Subject 5	545.9808	881.258
Subject 6	550.0676	1141.3383
Subject 7	542.7533	1135.5232
Subject 8	538.7697	1114.9791
Subject 9	541.0152	887.6454
Subject 10	534.8294	1111.1364

Table 6.6: Inference timings for 10 trained models of MI-classification (Left hand / Right hand / Feet) from the leave-one-subject-out cross-validation for EEG-Inception in PhysioNet. Measured with $torch.autograd.profiler$ in 2.9 GHz 6-core Central Processing Unit (CPU) Intel Core i9.

In contrast to other techniques that promise to tackle the issue of inter-subject variability, this framework is dynamically adapted to a new subject during inference without the need of re-training or calibration trials, commonly used in transfer learning methods. Furthermore, an inherent advantage of our framework is the training of K parallel personalised models of the same BCI architecture. During training, these models are not trained using only the samples of one specific subject but also samples from "similar" subjects since the attention mechanism is trained simultaneously. An interesting experiment would be to evaluate the performance of these inherent personalised models compared to standard personalised models - trained using strictly the samples of one specific subject. Although the BCI deep architectures ([62], [63], [64]) used this experimental section are considered state-of-the-art and achieve high performance across different MI-BCI tasks, they are usually comprised of thousands of trainable parameters, making the training of standard personalised models difficult with these publicly available datasets.

6.3.7 Ablation Study - Attention Comparison

As described previously in this chapter, during inference when an input EEG signal from a new unseen subject S_x is processed, it passes firstly through the attention network and the subject attention vector $\boldsymbol{\pi}$ is computed. Through investigation, we observed that this vector is quite sparse. Although this is generally something desirable, the low SNR of the EEG signal makes the dynamic framework unstable especially when used in the desired zero-calibration one-trial setup. In order to have a robust network that dynamically adapts to the new trial from an unseen subject, the "uniformly attended vector" $\mathbf{A^*}$ is utilised that uses knowledge from all K individuals and "shift" the attention more to the most relevant subjects. A comparison between using the vector $\boldsymbol{\pi}$ versus our uniformly attended vector $\mathbf{A^*}$ as attention in our dynamic framework can be seen in the following Figure 6.4.

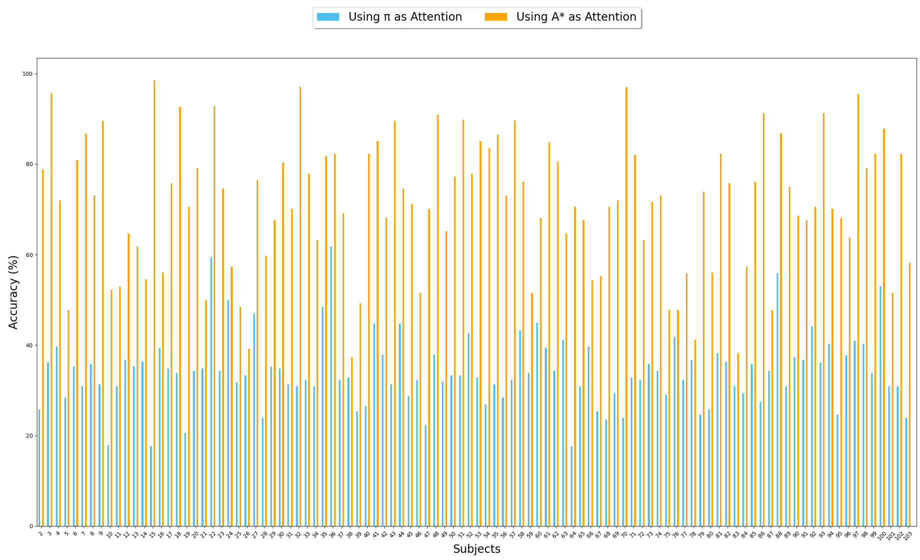

Figure 6.4: Comparison between using the vector $\boldsymbol{\pi}$ versus our uniformly attended vector $\mathbf{A^*}$ as attention in our dynamic framework. The performances correspond to generic (trained and evaluated in a leave-one-subject-out fashion) Dynamic EEG-Inception models in the MI-classification (Left / Right hand / Feet) task in PhysioNet.

6.4 Conclusions

In this chapter, we use the analysis (Figure 4.4) of the task of MI EEG classification through the lens of Causal Reasoning as described in Chapter 4. Through this analysis, we identify and analyze some of the major challenges and we introduce a framework based on dynamic convolutions that tackles the problem of subject distribution shift (inter-subject variability). Our subject attention mechanism achieves great performance in identifying subjects and the overall dynamic framework demonstrates increased performance when applied to different BCI architectures while at the same time outperforming other similar methods.

7
Conclusions

Contents

In this Thesis, we presented a comprehensive exploration of key challenges in BCI research and proposed several solutions, both theoretical and practical, to advance the field, ultimately paving the way for more effective and robust Brain-Computer Interfaces.

In Chapter 3, we introduced our EEG data collection experiment conducted at Imperial College London, aimed at capturing and analyzing real and imagined movements of participants engaged in diverse lab-structured and gaming scenarios. Through meticulous planning and execution, we orchestrated a comprehensive data collection experiment (that lasted more than two years) involving participants interacting with various gaming environments. By utilizing state-of-the-art hardware and EEG headsets, we successfully collected data samples corresponding to both physical actions and mental simulations from a significant number of participants.

In Chapter 4, we explored brainwave decoding using Causal Reasoning, noting the limitations of machine learning models in real-world BCI applications. We introduced a causal framework integrating Causal Reasoning into brainwave modeling to analyze challenges across the machine learning pipeline. We discussed how general machine learning practices and brainwave-specific techniques could address these challenges and proposed evaluation methods for measuring their performance.

Chapter 5 introduced a novel neural network architecture, leveraging Gabor filters to distribute EEG signal information across frequency and temporal modulations. The network showed high performance in both generic and personalized classification tasks, with fewer parameters and shorter training times compared to existing BCI architectures. Additionally, interpretability was enhanced, facilitating efficient personalized BCI models even with limited training data. Furthermore, a new subject selection framework combining discriminative learning and graph-signal processing was introduced, leading to improved personalized performance, especially for subjects with initially lower performance.

In Chapter 6, we tackled the issue of inter-subject variability in MI decoding, which hinders the generalization of deep models across subjects due to differing neurophysiological processes. To overcome this challenge, we introduced a dynamic convolution framework to account for shifts caused by inter-subject variability. Through experiments, we showed improved generalization performance across subjects in various MI tasks across different deep BCI architectures.

7.1 Future Work

The improvements and frameworks introduced in this Thesis open up new research directions that can be pursued to further build deep learning techniques for EEG-based Brain-Computer Interfaces.

The MyBrainCommands dataset presented in Chapter 3 fulfills the requirement for the need for EEG data that closely mirror real-world conditions beyond the confines of the laboratory. Although only a subset of this dataset was utilized in this Thesis, the whole dataset (with all its various settings) is currently being explored by the author as well as by his collaborators and deep learning models are being developed for real-life BCI applications [20] [21].

The innovative causal framework introduced in Chapter 4 marks a significant advancement, poised to serve as a cornerstone for the development of sophisticated BCI brainwave deep learning decoders. Drawing parallels to the field of Medical Imaging [95], we anticipate that this framework will establish itself as a guiding principle for future research in brainwave modeling. While initially applied in this Thesis (in Chapters 5 and 6) within the domain of MI-BCIs, researchers within the BCI community have already begun adopting it [198], and we are confident that its integration will persist across all BCI paradigms (beyond MI-BCIs).

[20] https://www.bbc.co.uk/news/technology-64720533

[21] https://news.sky.com/video/elon-musks-neuralink-company-implants-brain-chip-in-human-for-first-time-13060639

Our novel BrainWave-Scattering Net (presented Chapter 5) has proved itself, as demonstrated in other works [199], as a lightweight benchmark deep network with competitive generalized and personalised performances. Despite its performance as well as its advantages in transferability and interpreability properties, it contains certain biases. More specifically, it heavily relies on specific fixed kernel sizes of the underlying convolutional neural network architecture (e.g. the first convolutional layer has size of $(1, \frac{F_s}{2})$). Although these biases prohibit easier generalization across BCI datasets, where the sampling frequencies vary, they are not met only in our architecture but can also be found in all state-of-the-art deep learning BCI architectures (e.g. [62], [63], [64]) that this network was compared to and has demonstrated the benefits described in Chapter 5. An important factor specific to the network is the appropriate initialization values for the bandwidths and frequencies of the Gabor filters. As demonstrated in Chapter 5, these task-related inductive biases along with the trainable nature of the architecture can demonstrate a series of advantages and enhanced properties. Learning though these crucial inductive biases (kernel sizes and initialization values) directly from data would enable even greater generalization across BCI paradigms and tasks with various sampling rates and frequency contents while inheriting all its advantageous properties.

Guided by our causal framework, the dynamic convolution framework detailed in Chapter 6 has exhibited enhanced generalization performance across subjects in diverse MI tasks across a range of deep BCI architectures, establishing itself as a benchmark for comparison in the literature [200]. However, one drawback of this dynamic framework is the inevitable escalation in the number of trainable parameters (approximately $\times$ K, where K represents the number of subjects in the training dataset). Inspired by analogous studies [201], future endeavors could explore strategies to alleviate this issue.

Lastly, the next pivotal step in the author's research will be in the area of Generative AI, (generative brainwave models as described in Chapter 4), and more specifically the development of an efficient Brainwave Foundation Model (FM). In the last decade, advances in Deep Learning have enabled the development of powerful foundation model architectures that show remarkable success in areas like Natural Language Processing (NLP) and Computer Vision (e.g. [202], [203], [204], [205], [206]). Foundation Models, characterized by their pre-training on vast and varied data in a task-agnostic manner, typically through self-supervision, possess the readiness and adaptability for fine-tuning across a diverse array of downstream tasks. This obviates the necessity for large annotated datasets typically associated with supervised learning. The groundbreaking impact of these leading foundation model architectures spans across application domains, owing to their versatility, scalability, and ability in knowledge transfer without extensive task-specific customization (task-specific re-training). Additionally, their capability to generate highly realistic data further enhances their utility and applicability.

Within the BCI domain, foundation model architectures wield significant potential for EEG decoding. Initial efforts have already emerged in the BCI research community (e.g. [207], [208], [209], [210]). Their extensive self-supervised pre-training on diverse, unlabeled observations enables the capture of intricate patterns and relationships within EEG data. This augmentation notably enhances generalizability, allowing seamless adaptation to unseen tasks without necessitating full re-training, achieved through simple adaptable layers learned from a small amount of target data. Moreover, the transition of BCIs from laboratory settings to everyday user environments demands deep learning models adaptable to diverse consumer-grade devices, with resilience to signal irregularities. Brainwave Foundation Models hold promise in forecasting brain activities, reconstructing missing or damaged segments of brain signals, and refining artifact or noise removal from raw EEG signals. Such advancements strengthen the robustness of brainwave decoding models while facilitating the development of fully device-agnostic models - models which are deployable out-of-the-box across headsets featuring varied sensor configurations.

Bibliography

[1] A. A. Ninet, "Protecting the "homo digitalis"," *NAVEIÑ REET: Nordic Journal of Law and Social Research*, no. 9, pp. 153–170, Dec. 2019, ISSN: 2246-7483.

[2] G. Moore, "Cramming more components onto integrated circuits," *Proceedings of the IEEE*, vol. 86, no. 1, pp. 82–85, 1998.

[3] A. M. TURING, "COMPUTING MACHINERY AND INTELLIGENCE," *Mind*, vol. LIX, no. 236, pp. 433–460, Oct. 1950, ISSN: 0026-4423.

[4] Marcus, S. and Charles A. Dana Foundation, *Neuroethics: Mapping the Field : Conference Proceedings, May 13-14, 2002, San Francisco, California*. Dana Press, 2002, ISBN: 9780972383004.

[5] M. Ienca, "Common human rights challenges raised by different applications of neurotechnologies in the biomedical field," *Report commissioned by the Council of Europe*, 2021.

[6] S. UNESCO Milan-Bicocca University, *The Risks and Challenges of Neurotechnologies for Human Rights*. UNESCO Publishing, 2023, ISBN: 9789231005671.

[7] A. McCay, "Neurorights: The Chilean constitutional change," *AI & SOCIETY*, vol. 39, no. 2, pp. 797–798, 2024.

[8] K. Barmpas, Y. Panagakis, G. Zoumpourlis, D. A. Adamos, N. Laskaris, and S. Zafeiriou, "A causal perspective on brainwave modeling for brain–computer interfaces," *Journal of Neural Engineering*, vol. 21, no. 3, p. 036 001, May 2024.

[9] K. Barmpas, Y. Panagakis, D. A. Adamos, N. Laskaris, and S. Zafeiriou, "Brainwave-scattering net: A lightweight network for eeg-based motor imagery recognition," *Journal of Neural Engineering*, vol. 20, no. 5, p. 056 014, Sep. 2023, ISSN: 1741-2552.

[10] K. Barmpas, Y. Panagakis, D. Adamos, N. Laskaris, and S. Zafeiriou, "Subject selection framework to improve personalised models for motor-imagery BCIs via wavelets and graph diffusion," in *ICLR 2024 Workshop on Learning from Time Series For Health*, 2024.

[11] K. Barmpas, Y. Panagakis, S. Bakas, D. A. Adamos, N. Laskaris, and S. Zafeiriou, “Improving generalization of cnn-based motor-imagery eeg decoders via dynamic convolutions,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1997–2005, 2023.

[12] K. Barmpas, Y. Panagakis, D. A. Adamos, N. Laskaris, and S. Zafeiriou, “A CAUSAL VIEWPOINT ON MOTOR-IMAGERY BRAINWAVE DECODING,” *ICLR2022 Workshop on the Elements of Reasoning: Objects, Structure and Causality*, 2022.

[13] K. Barmpas, Y. Panagakis, D. A. Adamos, N. Laskaris, and S. Zafeiriou, “Classification of brain activity signals,” UK Patent GB2605270, Available online: https://www.ipo.gov.uk/p-ipsum/Case/PublicationNumber/GB2605270, 2024.

[14] K. Barmpas, Y. Panagakis, D. A. Adamos, N. Laskaris, and S. Zafeiriou, “Classification of brain activity signals,” European Patent EP4452055, Available online: https://register.epo.org/application?number=EP23702880, 2024.

[15] K. Barmpas, Y. Panagakis, D. A. Adamos, N. Laskaris, and S. Zafeiriou, “Classification of brain activity signals,” WPO PCT WO2023148471, 2023.

[16] J. Suk, L. Giusti, T. Hemo, M. Lopez, K. Barmpas, and C. Bodnar, “Surfing on the neural sheaf,” in *NeurIPS 2022 Workshop on Symmetry and Geometry in Neural Representations*, 2022.

[17] X. Wei, A. A. Faisal, M. Grosse-Wentrup, *et al.*, “2021 beetl competition: Advancing transfer learning for subject independence heterogenous eeg data sets,” *Proceedings of Machine Learning Research (PMLR)*, 2022.

[18] S. Bakas, S. Ludwig, K. Barmpas, *et al.*, *Team cogitat at neurips 2021: Benchmarks for eeg transfer learning competition*, 2022. arXiv: 2202.03267.

[19] M. Tudor, L. Tudor Car, and K. Tudor, “Hans berger (1873-1941) - the history of electroencephalography,” *Acta medica Croatica : časopis Hravatske akademije medicinskih znanosti*, vol. 59, pp. 307–13, Feb. 2005.

[20] H. Berger, “Über das elektrenkephalogramm des menschen,” *Archiv für Psychiatrie und Nervenkrankheiten*, vol. 87, no. 1, pp. 527–570, Dec. 1929, ISSN: 1433-8491.

[21] D. L. Schomer and F. H. Lopes da Silva, *Niedermeyer’s Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Oxford University Press, Nov. 2017, ISBN: 9780190228484.

[22] C. E. Stafstrom and L. Carmant, "Seizures and epilepsy: An overview for neuroscientists," *Cold Spring Harbor Perspectives in Medicine*, vol. 5, no. 6, a022426–a022426, Jun. 2015, ISSN: 2157-1422.

[23] D. E. Tan, R. S. Tung, W. Y. Leong, and J. C. M. Than, "Sleep disorder detection and identification," *Procedia Engineering*, vol. 41, pp. 289–295, 2012, ISSN: 1877-7058.

[24] A. T. Alouani and T. Elfouly, "Traumatic brain injury (tbi) detection: Past, present, and future," *Biomedicines*, vol. 10, no. 10, p. 2472, Oct. 2022, ISSN: 2227-9059.

[25] J. Kamiya, "Autoregulation of the eeg alpha rhythm: A program for the study of consciousness," in *Mind/Body Integration: Essential Readings in Biofeedback*. Boston, MA: Springer US, 1979, pp. 289–297, ISBN: 978-1-4613-2898-8.

[26] J. J. Vidal, "Toward direct brain-computer communication," *Annual Review of Biophysics and Bioengineering*, vol. 2, no. 1, pp. 157–180, Jun. 1973, ISSN: 0084-6589.

[27] L. Farwell and E. Donchin, "Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, Dec. 1988, ISSN: 0013-4694.

[28] N. Sobell and M. Trivich, "Brainwave drawing game," in *Delicate Balance: Technics, Culture and Consequences*, 1989, pp. 360–362.

[29] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An eeg-based brain-computer interface for cursor control," *Electroencephalography and Clinical Neurophysiology*, vol. 78, no. 3, pp. 252–259, Mar. 1991, ISSN: 0013-4694.

[30] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer, "Eeg-based discrimination between imagination of right and left hand movement," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 6, pp. 642–651, Dec. 1997, ISSN: 0013-4694.

[31] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial eeg during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.

[32] Schalk, "Can electrocorticography (ecog) support robust and powerful brain-computer interfaces?" *Frontiers in Neuroengineering*, 2010, ISSN: 1662-6443.

[33] C. Nam, A. Nijholt, and F. Lotte, *Brain–Computer Interfaces Handbook: Technological and Theoretical Advances*. CRC Press, 2018.

[34] M. Ferrari and V. Quaresima, "A brief review on the history of human functional near-infrared spectroscopy (fnirs) development and fields of application," *NeuroImage*, vol. 63, no. 2, pp. 921–935, Nov. 2012, ISSN: 1053-8119.

[35] G. H. Glover, "Overview of functional magnetic resonance imaging," *Neurosurgery Clinics of North America*, vol. 22, no. 2, pp. 133–139, Apr. 2011, ISSN: 1042-3680.

[36] H.-J. Hwang, S. Kim, S. Choi, and C.-H. Im, "Eeg-based brain-computer interfaces: A thorough literature survey," *International Journal of Human-Computer Interaction*, vol. 29, no. 12, pp. 814–826, Dec. 2013, ISSN: 1532-7590.

[37] G. H. Klem, H. Lüders, H. H. Jasper, and C. E. Elger, "The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology.," *Electroencephalography and clinical neurophysiology. Supplement*, vol. 52, pp. 3–6, 1999.

[38] P. K. Douglas and D. B. Douglas, "Reconsidering spatial priors in eeg source estimation : Does white matter contribute to eeg rhythms?" *2019 7th International Winter Conference on Brain-Computer Interface (BCI)*, pp. 1–12, 2019.

[39] M. X. Cohen, *Analyzing Neural Time Series Data: Theory and Practice*. The MIT Press, 2014, ISBN: 9780262319553.

[40] M. Chaumon, D. V. Bishop, and N. A. Busch, "A practical guide to the selection of independent components of the electroencephalogram for artifact correction," *Journal of Neuroscience Methods*, vol. 250, pp. 47–63, Jul. 2015, ISSN: 0165-0270.

[41] S. Debener, J. Thorne, T. R. Schneider, and F. C. Viola, "Using ica for the analysis of multichannel eeg data," in *Simultaneous EEG and fMRI*. Oxford University PressNew York, Apr. 2010, pp. 121–134, ISBN: 9780199776283.

[42] I. Winkler, S. Haufe, and M. Tangermann, "Automatic classification of artifactual ica-components for artifact removal in eeg signals," *Behavioral and Brain Functions*, vol. 7, no. 1, p. 30, 2011.

[43] M. J. Alhaddad, "Common average reference (car) improves p300 speller," 2012.

[44] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-r. Muller, "Optimizing spatial filters for robust eeg single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41–56, 2008.

[45] J. Decety and D. H. Ingvar, "Brain structures participating in mental simulation of motor behavior: A neuropsychological interpretation," *Acta Psychologica*, vol. 73, no. 1, pp. 13–34, 1990, ISSN: 0001-6918.

[46] M. Hallett, J. Fieldman, L. G. Cohen, N. Sadato, and A. Pascual-Leone, "Involvement of primary motor cortex in motor imagery and mental practice," *Behavioral and Brain Sciences*, vol. 17, no. 2, pp. 210–210, Jun. 1994, ISSN: 1469-1825.

[47] M. Jeannerod, "Neural simulation of action: A unifying mechanism for motor cognition," *NeuroImage*, vol. 14, no. 1, S103–S109, Jul. 2001, ISSN: 1053-8119.

[48] T. J. Kimberley, G. Khandekar, L. L. Skraba, J. A. Spencer, E. A. Van Gorp, and S. R. Walker, "Neural substrates for motor imagery in severe hemiparesis," *Neurorehabilitation and Neural Repair*, vol. 20, no. 2, pp. 268–277, Jun. 2006, ISSN: 1552-6844.

[49] R. H. Van der Lubbe, J. Sobierajewicz, M. L. Jongsma, W. B. Verwey, and A. Przekoracka-Krawczyk, "Frontal brain areas are more involved during motor imagery than during motor execution/preparation of a response sequence," *International Journal of Psychophysiology*, vol. 164, pp. 71–86, Jun. 2021, ISSN: 0167-8760.

[50] R. Mane, T. Chouhan, and C. Guan, "BCI for stroke rehabilitation: Motor and beyond," *Journal of Neural Engineering*, vol. 17, no. 4, p. 041 001, Aug. 2020.

[51] N. Robinson, R. Mane, T. Chouhan, and C. Guan, "Emerging trends in bci-robotics for motor control and rehabilitation," *Current Opinion in Biomedical Engineering*, vol. 20, p. 100 354, 2021, ISSN: 2468-4511.

[52] M. Sebastián-Romagosa, W. Cho, R. Ortner, *et al.*, "Brain computer interface treatment for motor rehabilitation of upper extremity of stroke patients—a feasibility study," *Frontiers in Neuroscience*, vol. 14, p. 1056, 2020, ISSN: 1662-453X.

[53] J. Decety and D. H. Ingvar, "Brain structures participating in mental simulation of motor behavior: A neuropsychological interpretation," *Acta Psychologica*, vol. 73, no. 1, pp. 13–34, 1990, ISSN: 0001-6918.

[54] S. Saha and M. Baumert, "Intra- and inter-subject variability in eeg-based sensorimotor brain computer interface: A review," *Frontiers in Computational Neuroscience*, vol. 13, p. 87, ISSN: 1662-5188.

[55] B. Guragai, O. AlShorman, M. Masadeh, and M. B. B. Heyat, "A survey on deep learning classification algorithms for motor imagery," in *2020 32nd International Conference on Microelectronics (ICM)*, 2020, pp. 1–4.

[56] D. McFarland, C. Anderson, K.-R. Muller, A. Schlogl, and D. Krusienski, "Bci meeting 2005-workshop on bci signal processing: Feature extraction and translation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 135–138, 2006.

[57] F. Lotte and C. Guan, “Regularizing common spatial patterns to improve bci designs: Unified theory and new algorithms,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 355–362, 2011.

[58] W. Wu, Z. Chen, X. Gao, Y. Li, E. N. Brown, and S. Gao, “Probabilistic common spatial patterns for multichannel eeg analysis,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 639–653, 2015.

[59] E. A. Mousavi, J. J. Maller, P. B. Fitzgerald, and B. J. Lithgow, “Wavelet common spatial pattern in asynchronous offline brain computer interfaces,” *Biomedical Signal Processing and Control*, vol. 6, no. 2, pp. 121–128, 2011, Special Issue: The Advance of Signal Processing for Bioelectronics, ISSN: 1746-8094.

[60] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, “Filter bank common spatial pattern (fbcsp) in brain-computer interface,” in *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 2008, pp. 2390–2397.

[61] Y. Miao, J. Jin, I. Daly, *et al.*, “Learning common time-frequency-spatial patterns for motor imagery classification,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 699–707, 2021.

[62] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, *et al.*, “Deep learning with convolutional neural networks for eeg decoding and visualization,” *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[63] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, “EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces,” *Journal of Neural Engineering*, vol. 15, no. 5, p. 056 013, Jul. 2018.

[64] E. Santamaría-Vázquez, V. Martínez-Cagigal, F. Vaquerizo-Villar, and R. Hornero, “Eeg-inception: A novel deep convolutional neural network for assistive erp-based brain-computer interfaces,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 12, pp. 2773–2782, 2020.

[65] C. Szegedy, W. Liu, Y. Jia, *et al.*, “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.

[66] A. Goldberger, L. Amaral, L. Glass, *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220.,” *Publicly Accessible Dataset*, 2000. [Online]. Available: `https://physionet.org/content/eegmmidb/1.0.0/`.

[67] K.-R. Müller, B. Blankertz, and G. Curio, "Data set IVa ‹motor imagery, small training sets›," *Publicly Accessible Dataset*, 2004. [Online]. Available: `http://www.bbci.de/competition/iii/desc_IVa.html`.

[68] O.-Y. Kwon, M.-H. Lee, C. Guan, and S.-W. Lee, "Subject-independent brain–computer interfaces based on deep convolutional neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 10, pp. 3839–3852, 2020. [Online]. Available: `https://github.com/PatternRecognition/OpenBMI`.

[69] N. P. Castellanos and V. A. Makarov, "Recovering eeg brain signals: Artifact suppression with wavelet enhanced independent component analysis," *Journal of Neuroscience Methods*, vol. 158, no. 2, pp. 300–312, 2006, ISSN: 0165-0270.

[70] S. Bakas, D. A. Adamos, and N. Laskaris, "On the estimate of music appraisal from surface EEG: A dynamic-network approach based on cross-sensor PAC measurements," *Journal of Neural Engineering*, vol. 18, no. 4, p. 046 073, Jun. 2021.

[71] C.-H. Han, K.-R. Müller, and H.-J. Hwang, "Brain-switches for asynchronous brain–computer interfaces: A systematic review," *Electronics*, vol. 9, no. 3, 2020, ISSN: 2079-9292.

[72] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain–computer interfaces for communication and rehabilitation," *Nat Rev Neurol*, vol. 12, 2016.

[73] T. Luu, S. Nakagome, Y. He, and J. Contreras-Vidal, "Real-time eeg-based brain-computer interface to a virtual avatar enhances cortical involvement in human," *Sci Rep*, vol. 7, 2017.

[74] G. Sharma, D. Friedenberg, N. Annetta, *et al.*, "Using an artificial neural bypass to restore cortical control of rhythmic movements in a human with quadriplegia," *Sci Rep*, vol. 6, 2016.

[75] A. Biasiucci, R. Leeb, I. Iturrate, *et al.*, "Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke.," *Nat Commun*, vol. 9, 2018.

[76] A. Bashashati, M. Fatourechi, R. K. Ward, and G. E. Birch, "A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals," *Journal of Neural Engineering*, vol. 4, no. 2, R32–R57, Mar. 2007.

[77] T. C. Handy, *Brain signal analysis advances in neuroelectric and neuromagnetic methods. Cambridge, Mass, MIT Press.* 2009.

[78] R. P. N. Rao, *Brain-computer interfacing: an introduction.* 2013.

[79] C. S. Nam, A. Nijholt, and F. Lotte, *Brain–Computer Interfaces Handbook: Technological and Theoretical Advances, CRC Press.* 2018.

[80] F. Lotte, L. Bougrain, A. Cichocki, *et al.*, "A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update," *Journal of Neural Engineering*, vol. 15, no. 3, p. 031 005, Apr. 2018.

[81] J. Thomas, T. Maszczyk, N. Sinha, T. Kluge, and J. Dauwels, "Deep learning-based classification for brain-computer interfaces," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 234–239.

[82] X. Zhang, L. Yao, X. Wang, J. Monaghan, D. McAlpine, and Y. Zhang, "A survey on deep learning-based non-invasive brain signals: Recent advances and new frontiers," *Journal of Neural Engineering*, vol. 18, no. 3, p. 031 002, Mar. 2021.

[83] A. Esteva, K. Chou, S. Yeung, *et al.*, "Deep learning-enabled medical computer vision," *npj Digital Medicine*, vol. 4, no. 1, p. 5, Jan. 2021, ISSN: 2398-6352.

[84] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, Y. Bengio and Y. LeCun, Eds., 2015.

[85] H. Purwins, B. Li, T. Virtanen, J. Schluter, S.-Y. Chang, and T. Sainath, "Deep learning for audio signal processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 2, pp. 206–219, May 2019, ISSN: 1941-0484.

[86] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, Minneapolis, Minnesota: Association for Computational Linguistics, Jun. 2019, pp. 4171–4186.

[87] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, "ALBERT: A lite BERT for self-supervised learning of language representations," *CoRR*, vol. abs/1909.11942, 2019. arXiv: `1909.11942`.

[88] A. Antoniades, L. Spyrou, C. C. Took, and S. Sanei, "Deep learning for epileptic intracranial eeg data," in *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2016, pp. 1–6.

[89] M.-P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, "Optimized deep learning for eeg big data and seizure prediction bci via internet of things," *IEEE Transactions on Big Data*, vol. 3, no. 4, pp. 392–404, 2017.

[90] M. Längkvist, L. Karlsson, and A. Loutfi, “Sleep stage classification using unsupervised feature learning,” *Advances in Artificial Neural Systems*, vol. 2012, Jul. 2012.

[91] D. F. Wulsin, J. R. Gupta, R. Mani, J. A. Blanco, and B. Litt, “Modeling electroencephalography waveforms with semi-supervised deep belief nets: Fast classification and anomaly measurement,” vol. 8, no. 3, p. 036 015, Apr. 2011.

[92] K. Kumarasinghe, N. Kasabov, and D. Taylor, “Brain-inspired spiking neural networks for decoding and understanding muscle activity and kinematics from electroencephalography signals during hand movements,” *Sci Rep*, vol. 11, 2021.

[93] L. D. Sharma, V. K. Bohat, M. Habib, A. M. Al-Zoubi, H. Faris, and I. Aljarah, “Evolutionary inspired approach for mental stress detection using eeg signal,” *Expert Systems with Applications*, vol. 197, p. 116 634, Jul. 2022, ISSN: 0957-4174.

[94] Y. Rezaeitabar and U. Halici, “A novel deep learning approach for classification of eeg motor imagery signals,” *Journal of Neural Engineering*, vol. 14, p. 016 003, Feb. 2017.

[95] D. Castro, I. Walker, and B. Glocker, “Causality matters in medical imaging. Commun 11, 3673,” 2020.

[96] B. Schölkopf, F. Locatello, S. Bauer, *et al.*, “Toward causal representation learning,” *Proceedings of the IEEE*, vol. 109, no. 5, pp. 612–634, 2021.

[97] I. J. Goodfellow, J. Shlens, and C. Szegedy, *Explaining and harnessing adversarial examples*, 2015. arXiv: `1412.6572`.

[98] B. Schölkopf, D. Janzing, J. Peters, E. Sgouritsa, K. Zhang, and J. M. Mooij, “On causal and anticausal learning,” in *ICML*, 2012.

[99] J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference—Foundations and Learning Algorithms*. Cambridge, MA, USA: MIT Press, 2017.

[100] H. Reichenbach, *The Direction of Time*. Dover Publications, 1956.

[101] P. Ritter, M. Schirner, A. R. McIntosh, and V. K. Jirsa, “The virtual brain integrates computational modeling and multimodal neuroimaging,” *Brain Connectivity*, vol. 3, no. 2, pp. 121–145, 2013, PMID: 23442172.

[102] K. K. Ang, C. Guan, K. S. Phua, *et al.*, “Transcranial direct current stimulation and eeg-based motor imagery bci for upper limb stroke rehabilitation,” in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 4128–4131.

[103] M. Beudel and P. Brown, "Adaptive deep brain stimulation in parkinson's disease," *Parkinsonism & Related Disorders*, vol. 22, S123–S126, 2016, Proceedings of XXI World Congress on Parkinson's Disease and Related Disorders, December 6-9, 2015, Milan, Italy, ISSN: 1353-8020.

[104] S. Weichwald, T. Meyer, M. Grosse-Wentrup, and B. Schölkopf, "Causal interpretation rules for encoding and decoding models in neuroimaging," Jan. 2015.

[105] K. Georgiadis, D. A. Adamos, S. Nikolopoulos, N. Laskaris, and I. Kompatsiaris, "Covariation informed graph slepians for motor imagery decoding," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 340–349, 2021.

[106] T. O. Zander and C. Kothe, "Towards passive brain–computer interfaces: Applying brain–computer interface technology to human–machine systems in general," *Journal of Neural Engineering*, vol. 8, no. 2, p. 025 005, Mar. 2011.

[107] K. Won, M. Kwon, S. Jang, M. Ahn, and S. C. Jun, "P300 speller performance predictor based on rsvp multi-feature," *Frontiers in Human Neuroscience*, vol. 13, 2019, ISSN: 1662-5161.

[108] V. Abootalebi, M. H. Moradi, and M. A. Khalilzadeh, "A new approach for eeg feature extraction in p300-based lie detection," *Computer Methods and Programs in Biomedicine*, vol. 94, no. 1, pp. 48–57, 2009, ISSN: 0169-2607.

[109] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: A review," *Journal of vision*, vol. 15, no. 6, pp. 4–4, 2015.

[110] R. Chavarriaga and J. d. R. Millán, "Learning from eeg error-related potentials in noninvasive brain-computer interfaces," *IEEE transactions on neural systems and rehabilitation engineering*, vol. 18, no. 4, pp. 381–388, 2010.

[111] K. Liao, R. Xiao, J. Gonzalez, and L. Ding, "Decoding individual finger movements from one hand using human eeg signals," *PLOS ONE*, vol. 9, no. 1, e85192, 2014.

[112] M. V. Kosti, K. Georgiadis, D. A. Adamos, N. Laskaris, D. Spinellis, and L. Angelis, "Towards an affordable brain computer interface for the assessment of programmers' mental workload," *International Journal of Human-Computer Studies*, vol. 115, pp. 52–66, 2018, ISSN: 1071-5819.

[113] G. Bin, X. Gao, Y. Wang, Y. Li, B. Hong, and S. Gao, "A high-speed bci based on code modulation vep," *Journal of neural engineering*, vol. 8, no. 2, p. 025 015, 2011.

[114] B. Kaneshiro, M. Perreau Guimaraes, H.-S. Kim, A. M. Norcia, and P. Suppes, "A representational similarity analysis of the dynamics of object processing using single-trial eeg classification," *Plos one*, vol. 10, no. 8, e0135697, 2015.

[115] C. Sas and R. Chopra, "Meditaid: A wearable adaptive neurofeedback-based system for training mindfulness state," *Personal and Ubiquitous Computing*, vol. 19, pp. 1169–1182, 2015.

[116] H. Phan, O. Y. Chén, M. C. Tran, P. Koch, A. Mertins, and M. De Vos, "Xsleepnet: Multi-view sequential model for automatic sleep staging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5903–5915, 2021.

[117] J. A. O'sullivan, A. J. Power, N. Mesgarani, *et al.*, "Attentional selection in a cocktail party environment can be decoded from single-trial eeg," *Cerebral cortex*, vol. 25, no. 7, pp. 1697–1706, 2015.

[118] F. Kalaganis, D. Adamos, and N. Laskaris, "Musical neuropicks: A consumer-grade bci for on-demand music streaming services," *Neurocomputing*, vol. 280, pp. 65–75, 2018, Applications of Neural Modeling in the new era for data and IT, ISSN: 0925-2312.

[119] A. Ragazzoni, F. Di Russo, S. Fabbri, *et al.*, ""hit the missing stimulus". a simultaneous eeg-fmri study to localize the generators of endogenous erps in an omitted target paradigm," *Scientific reports*, vol. 9, no. 1, p. 3684, 2019.

[120] C.-T. Lin, C.-J. Chang, B.-S. Lin, S.-H. Hung, C.-F. Chao, and I.-J. Wang, "A real-time wireless brain–computer interface system for drowsiness detection," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 4, no. 4, pp. 214–222, 2010.

[121] S. Olsen, G. Alder, M. Williams, *et al.*, "Electroencephalographic recording of the movement-related cortical potential in ecologically valid movements: A scoping review," *Frontiers in Neuroscience*, vol. 15, 2021, ISSN: 1662-453X.

[122] C. He, J. Liu, Y. Zhu, and W. Du, "Data augmentation for deep neural networks model in eeg classification task: A review," *Frontiers in Human Neuroscience*, vol. 15, 2021, ISSN: 1662-5161.

[123] Y. Luo and B.-L. Lu, "Eeg data augmentation for emotion recognition using a conditional wasserstein gan," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 2535–2538.

[124] F. Fahimi, Z. Zhang, W. B. Goh, K. K. Ang, and C. Guan, "Towards eeg generation using gans for bci applications," in *2019 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, 2019, pp. 1–4.

[125] Y. Song, L. Yang, X. Jia, and L. Xie, "Common spatial generative adversarial networks based eeg data augmentation for cross-subject brain-computer interface," *ArXiv*, vol. abs/2102.04456, 2021.

[126] M. Parvan, A. R. Ghiasi, T. Y. Rezaii, and A. Farzamnia, "Transfer learning based motor imagery classification using convolutional neural networks," in *2019 27th Iranian Conference on Electrical Engineering (ICEE)*, 2019, pp. 1825–1828.

[127] Z. Yin and J. Zhang, "Cross-session classification of mental workload levels using eeg and an adaptive deep learning model," *Biomedical Signal Processing and Control*, vol. 33, pp. 30–47, 2017, ISSN: 1746-8094.

[128] N.-S. Kwak, K.-R. Müller, and S.-W. Lee, "A convolutional neural network for steady state visual evoked potential classification under ambulatory environment," *PLOS ONE*, vol. 12, pp. 1–20, Feb. 2017.

[129] G. Dai, J. Zhou, J. Huang, and N. Wang, "HS-CNN: A CNN with hybrid convolution scale for EEG motor imagery classification," *Journal of Neural Engineering*, vol. 17, no. 1, p. 016 025, Jan. 2020.

[130] F. P. Kalaganis, N. A. Laskaris, E. Chatzilari, S. Nikolopoulos, and I. Kompatsiaris, "A data augmentation scheme for geometric deep learning in personalized brain–computer interfaces," *IEEE Access*, vol. 8, pp. 162 218–162 229, 2020.

[131] N. K. Nik Aznan, A. Atapour-Abarghouei, S. Bonner, J. D. Connolly, N. Al Moubayed, and T. P. Breckon, "Simulating brain signals: Creating synthetic eeg data via neural-based generative models for improved ssvep classification," in *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019, pp. 1–8.

[132] J. Behncke, R. T. Schirrmeister, M. Völker, A. Schulze-Bonhage, W. Burgard, and T. Ball, "Cross-paradigm pretraining of convolutional networks improves intracranial EEG decoding," *CoRR*, vol. abs/1806.09532, 2018. arXiv: 1806.09532.

[133] M. T. Sadiq, M. Z. Aziz, A. Almogren, A. Yousaf, S. Siuly, and A. U. Rehman, "Exploiting pretrained cnn models for the development of an eeg-based robust bci framework," *Computers in Biology and Medicine*, vol. 143, p. 105 242, 2022, ISSN: 0010-4825.

[134] B. Scholkopf, *Causality for Machine Learning*. 2019.

[135] A. M. Abdelhameed and M. Bayoumi, “Semi-supervised eeg signals classification system for epileptic seizure detection,” *IEEE Signal Processing Letters*, vol. 26, no. 12, pp. 1922–1926, 2019.

[136] Z. Gu, Z. Yu, Z. Shen, and Y. Li, “An online semi-supervised brain–computer interface,” *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 9, pp. 2614–2623, 2013.

[137] D. Kostas, S. Aroca-Ouellette, and F. Rudzicz, “Bendr: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of eeg data,” *Frontiers in Human Neuroscience*, vol. 15, 2021, ISSN: 1662-5161.

[138] A. Gramfort, H. Banville, O. Chehab, A. Hyvärinen, and D. Engemann, “Learning with self-supervision on eeg data,” in *2021 9th International Winter Conference on Brain-Computer Interface (BCI)*, 2021, pp. 1–2.

[139] G. H. Klem, H. Lüders, H. H. Jasper, and C. E. Elger, “The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology.,” *Electroencephalography and clinical neurophysiology. Supplement*, vol. 52, pp. 3–6, 1999.

[140] S. Saha and M. Baumert, “Intra- and inter-subject variability in eeg-based sensorimotor brain computer interface: A review,” *Frontiers in Computational Neuroscience*, vol. 13, 2020, ISSN: 1662-5188.

[141] H. He and D. Wu, “Transfer learning for brain–computer interfaces: A euclidean space data alignment approach,” *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 399–410, 2020.

[142] O. Özdenizci, Y. Wang, T. Koike-Akino, and D. Erdoğmuş, “Learning invariant representations from eeg via adversarial inference,” *IEEE Access*, vol. 8, pp. 27 074–27 085, 2020.

[143] H. He and D. Wu, “Transfer learning for brain–computer interfaces: A euclidean space data alignment approach,” *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 399–410, 2019.

[144] T. Radüntz, J. Scouten, O. Hochmuth, and B. Meffert, “Automated EEG artifact elimination by applying machine learning algorithms to ICA-based features,” *Journal of Neural Engineering*, vol. 14, no. 4, p. 046 004, May 2017.

[145] P. K. Johal and N. Jain, “Artifact removal from eeg: A comparison of techniques,” in *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, 2016, pp. 2088–2091.

[146] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from eeg signals: A review," *Sensors*, vol. 19, no. 5, 2019, ISSN: 1424-8220.

[147] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal—state-of-the-art and guidelines," *Journal of Neural Engineering*, vol. 12, no. 3, p. 031 001, Apr. 2015.

[148] A. H. H. Al-nuaimi, E. Jammeh, L. Sun, and E. Ifeachor, "Complexity measures for quantifying changes in electroencephalogram in alzheimer's disease," *Complexity*, vol. 2018, pp. 1–12, Mar. 2018.

[149] B. Somers and A. Bertrand, "Removal of eye blink artifacts in wireless EEG sensor networks using reduced-bandwidth canonical correlation analysis," *Journal of Neural Engineering*, vol. 13, no. 6, p. 066 008, Oct. 2016.

[150] Z. WU and N. E. HUANG, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Advances in Adaptive Data Analysis*, vol. 01, no. 01, pp. 1–41, 2009.

[151] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals—practices and possibilities," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 3, pp. 488–500, 2012.

[152] D. Safieddine, A. Kachenoura, L. Albera, *et al.*, "Removal of muscle artifact from eeg data: Comparison between stochastic (ica and cca) and deterministic (emd and wavelet-based) approaches," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, Jul. 2012.

[153] Q. Yuan, W. Zhou, L. Zhang, *et al.*, "Epileptic seizure detection based on imbalanced classification and wavelet packet transform," *Seizure*, vol. 50, pp. 99–108, 2017, ISSN: 1059-1311.

[154] A. Barbu, D. Mayo, J. Alverio, *et al.*, "Objectnet: A large-scale bias-controlled dataset for pushing the limits of object recognition models," in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds., vol. 32, Curran Associates, Inc., 2019.

[155] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[156] F. Lotte, "Signal processing approaches to minimize or suppress calibration time in oscillatory activity-based brain–computer interfaces," *Proceedings of the IEEE*, vol. 103, no. 6, pp. 871–890, 2015.

[157] Z. Khademi, F. Ebrahimi, and H. M. Kordy, “A review of critical challenges in mi-bci: From conventional to deep learning methods,” *Journal of Neuroscience Methods*, vol. 383, p. 109 736, 2023, ISSN: 0165-0270.

[158] W. Xiong and Q. Wei, “Reducing calibration time in motor imagery-based bcis by data alignment and empirical mode decomposition,” *PLOS ONE*, vol. 17, no. 2, pp. 1–18, Feb. 2022.

[159] R. Scherer, A. Schloegl, F. Lee, H. Bischof, J. Janša, and G. Pfurtscheller, “The self-paced graz brain-computer interface: Methods and applications,” *Computational Intelligence and Neuroscience*, vol. 2007, 2007.

[160] S. Saha and M. Baumert, “Intra- and inter-subject variability in eeg-based sensorimotor brain computer interface: A review,” *Frontiers in Computational Neuroscience*, vol. 13, 2020, ISSN: 1662-5188.

[161] B. Blankertz, C. Sannelli, S. Halder, *et al.*, “Neurophysiological predictor of smr-based bci performance,” *NeuroImage*, vol. 51, no. 4, pp. 1303–1309, 2010, ISSN: 1053-8119.

[162] R. Zhang, F. Li, T. Zhang, D. Yao, and P. Xu, “Subject inefficiency phenomenon of motor imagery brain-computer interface: Influence factors and potential solutions.,” *Brain Science Advances*, vol. 6, no. 3, pp. 224–241, 2020.

[163] J. Anden, V. Lostanlen, and S. Mallat, “Joint time–frequency scattering,” *IEEE Transactions on Signal Processing*, vol. 67, no. 14, pp. 3704–3718, Jul. 2019, ISSN: 1941-0476.

[164] N. Zeghidour, O. Teboul, F. de Chaumont Quitry, and M. Tagliasacchi, “Leaf: A learnable frontend for audio classification,” in *International Conference on Learning Representations*, 2021.

[165] N. Zeghidour, N. Usunier, I. Kokkinos, T. Schaiz, G. Synnaeve, and E. Dupoux, “Learning filterbanks from raw speech for phone recognition,” in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 5509–5513.

[166] P.-G. Noé, T. Parcollet, and M. Morchid, “Cgcnn: Complex gabor convolutional neural network on raw speech,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 7724–7728.

[167] M. Ravanelli and Y. Bengio, “Speaker recognition from raw waveform with sincnet,” *arXiv preprint arXiv:1808.00158*, 2019. arXiv: `1808.00158`.

[168] R. Balestriero, R. Cosentino, H. Glotin, and R. Baraniuk, "Spline filters for end-to-end deep learning," in *Proceedings of the 35th International Conference on Machine Learning*, J. Dy and A. Krause, Eds., ser. Proceedings of Machine Learning Research, vol. 80, PMLR, Oct. 2018, pp. 364–373.

[169] S. Mallat, "Group invariant scattering," *Commun. Pure Appl. Math.*, no. 10, pp. 1331–1398, 2012.

[170] Y. Meyer, *Wavelets and Operators* (Cambridge Studies in Advanced Mathematics), D. H. Salinger, Ed. Cambridge University Press, 1993, vol. 1.

[171] I. Waldspurger, "Exponential decay of scattering coefficients," in *2017 International Conference on Sampling Theory and Applications (SampTA)*, 2017, pp. 143–146.

[172] J. Bruna and S. Mallat, "Classification with scattering operators," in *CVPR 2011*, 2011, pp. 1561–1566.

[173] B. Joan and M. Stephane, "Invariant scattering convolution networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1872–1886, 2013.

[174] S. Mallat, *A Wavelet Tour of Signal Processing : The Sparse Way. With contributions from Gabriel Peyré.* Academic Press, 3 edition, 2009.

[175] D. Gabor, English, *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, vol. 93, 429–441(12), 26 Nov. 1946, ISSN: 0367-7540.

[176] C. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, Jan. 1949, ISSN: 0096-8390. DOI: 10.1109/jrproc.1949.232969. [Online]. Available: http://dx.doi.org/10.1109/JRPROC.1949.232969.

[177] J. Kossaifi, A. Toisoul, A. Bulat, Y. Panagakis, T. M. Hospedales, and M. Pantic, "Factorized higher-order cnns with an application to spatio-temporal emotion estimation," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 6059–6068.

[178] Y. Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selection in motor imagery based brain-computer interface," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5392–5395, 2005.

[179] K. Georgiadis, D. A. Adamos, S. Nikolopoulos, N. Laskaris, and I. Kompatsiaris, "A graph-theoretic sensor-selection scheme for covariance-based motor imagery (mi) decoding," in *2020 28th European Signal Processing Conference (EUSIPCO)*, 2021, pp. 1234–1238.

[180] K. Georgiadis, N. Laskaris, S. Nikolopoulos, and I. Kompatsiaris, "Exploiting the heightened phase synchrony in patients with neuromuscular disease for the establishment of efficient motor imagery bcis," *Journal of neuroengineering and rehabilitation*, vol. 15, 2018.

[181] K. Georgiadis, D. Adamos, S. Nikolopoulos, N. Laskaris, and I. Kompatsiaris, "Covariation informed graph slepians for motor imagery decoding," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 340–349, 2021.

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

[183] M. Andreux, T. Angles, G. Exarchakis, *et al.*, "Kymatio: Scattering transforms in python," *Journal of Machine Learning Research*, vol. 21, no. 60, pp. 1–6, 2020.

[184] L. Sifre and J. Anden, "Scatnet," *Matlab Computer Software*, 2013. [Online]. Available: `http://www.di.ens.fr/data/software/scatnet`.

[185] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, no. 6, p. 80, Dec. 1945, ISSN: 0099-4987. DOI: `10.2307/3001968`. [Online]. Available: `http://dx.doi.org/10.2307/3001968`.

[186] N. Laskaris, D. Adamos, and A. Bezerianos, "Graph theory for brain signal processing," in *Handbook of Neuroengineering*, N. V. Thakor, Ed. Singapore: Springer Singapore, 2020, pp. 1–29, ISBN: 978-981-15-2848-4.

[187] N. Laskaris and A. Ioannides, "Semantic geodesic maps: A unifying geometrical approach for studying the structure and dynamics of single trial evoked responses," *Clinical Neurophysiology*, vol. 113, no. 8, pp. 1209–1226, 2002, ISSN: 1388-2457.

[188] S. Marcel and J. D. R. Millan, "Person authentication using brainwaves (eeg) and maximum a posteriori model adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, pp. 743–752, 2007.

[189] A. Valsaraj, I. Madala, N. Garg, M. Patil, and V. Baths, "Motor imagery based multimodal biometric user authentication system using eeg," in *2020 International Conference on Cyberworlds (CW)*, 2020, pp. 272–279.

[190] L. Yang, A. Libert, and M. M. Van Hulle, "Chronic study on brainwave authentication in a real-life setting: An lstm-based bagging approach," *Biosensors*, vol. 11, no. 10, 2021, ISSN: 2079-6374.

[191] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, “Multiclass brain–computer interface classification by riemannian geometry,” *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 920–928, 2012.

[192] H. Kang, Y. Nam, and S. Choi, “Composite common spatial pattern for subject-to-subject transfer,” *IEEE Signal Processing Letters*, vol. 16, pp. 683–686, 2009.

[193] D. Zhao, F. Tang, B. Si, and X. Feng, “Learning joint space–time–frequency features for eeg decoding on small labeled data,” *Neural Networks*, vol. 114, pp. 67–77, 2019, ISSN: 0893-6080.

[194] A. N. Olesen, P. Jennum, E. Mignot, and H. B. D. Sorensen, “Deep transfer learning for improving single-eeg arousal detection,” in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 99–103.

[195] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, “Adaptive transfer learning for eeg motor imagery classification with deep convolutional neural network,” *Neural Networks*, vol. 136, pp. 1–10, 2021, ISSN: 0893-6080.

[196] R. Zhang, Q. Zong, L. Dou, X. Zhao, Y. Tang, and Z. Li, “Hybrid deep neural network using transfer learning for eeg motor imagery decoding,” *Biomedical Signal Processing and Control*, vol. 63, p. 102 144, 2021, ISSN: 1746-8094.

[197] Y. Chen, X. Dai, M. Liu, D. Chen, L. Yuan, and Z. Liu, “Dynamic convolution: Attention over convolution kernels,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 11 027–11 036.

[198] X.-C. Zhong, Q. Wang, D. Liu, *et al.*, *Eeg-dg: A multi-source domain generalization framework for motor imagery eeg classification*, 2023. arXiv: 2311.05415.

[199] M. A. Abbasi, H. F. Abbasi, M. Z. Aziz, W. Haider, Z. Fan, and X. Yu, “A novel precisely designed compact convolutional eeg classifier for motor imagery classification,” *Signal, Image and Video Processing*, vol. 18, no. 4, pp. 3243–3254, Feb. 2024, ISSN: 1863-1711.

[200] G. Zoumpourlis and I. Patras, *Motor imagery decoding using ensemble curriculum learning and collaborative training*, 2024. arXiv: 2211.11460.

[201] F. Wu, A. Fan, A. Baevski, Y. Dauphin, and M. Auli, “Pay less attention with lightweight and dynamic convolutions,” in *International Conference on Learning Representations*, 2019.

[202] T. B. Brown, B. Mann, N. Ryder, *et al.*, *Language models are few-shot learners*, 2020. arXiv: 2005.14165.

[203] H. Touvron, T. Lavril, G. Izacard, *et al.*, *Llama: Open and efficient foundation language models*, 2023. arXiv: 2302.13971.

[204] L. Ouyang, J. Wu, X. Jiang, *et al.*, "Training language models to follow instructions with human feedback," in *Advances in Neural Information Processing Systems*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, Eds., vol. 35, Curran Associates, Inc., 2022, pp. 27 730–27 744.

[205] J. Wei, Y. Tay, R. Bommasani, *et al.*, *Emergent abilities of large language models*, 2022. arXiv: 2206.07682.

[206] F. P. Papantoniou, A. Lattas, S. Moschoglou, J. Deng, B. Kainz, and S. Zafeiriou, *Arc2face: A foundation model of human faces*, 2024. arXiv: 2403.11641.

[207] W. Jiang, L. Zhao, and B.-l. Lu, "Large brain model for learning generic representations with tremendous EEG data in BCI," in *The Twelfth International Conference on Learning Representations*, 2024.

[208] W. Cui, W. Jeong, P. Thölke, *et al.*, *Neuro-gpt: Towards a foundation model for eeg*, 2024. arXiv: 2311.03764 [cs.LG].

[209] K. Barmpas, Y. Panagakis, D. Adamos, N. Laskaris, and S. Zafeiriou, "A causal perspective in brainwave foundation models," in *Causality and Large Models @NeurIPS 2024*, 2024.

[210] K. Barmpas, G. Zoumpourlis, Y. Panagakis, D. Adamos, N. Laskaris, and S. Zafeiriou, "Position: Addressing ethical challenges and safety risks in genAI-powered brain-computer interfaces," in *NeurIPS 2024 GenAI for Health: Potential, Trust and Policy Compliance*, 2024.

A

Detailed Personalised Performances of BrainWave-Scattering Net

Subject	Task	Accuracy
Subject 1	MI Left vs Right Hand	68.8%
Subject 2	MI Left vs Right Hand	60%
Subject 3	MI Left vs Right Hand	55.5%
Subject 4	MI Left vs Right Hand	55.5%
Subject 5	MI Left vs Right Hand	60%
Subject 6	MI Left vs Right Hand	57.7%
Subject 7	MI Left vs Right Hand	91.1%
Subject 8	MI Left vs Right Hand	51.1%
Subject 9	MI Left vs Right Hand	55.5%
Subject 10	MI Left vs Right Hand	66.6%
Subject 11	MI Left vs Right Hand	57.7%
Subject 12	MI Left vs Right Hand	48.8%
Subject 13	MI Left vs Right Hand	48.8%
Subject 14	MI Left vs Right Hand	57.7%
Subject 15	MI Left vs Right Hand	75.5%
Subject 16	MI Left vs Right Hand	57.7%
Subject 17	MI Left vs Right Hand	64.4%
Subject 18	MI Left vs Right Hand	53.3%
Subject 19	MI Left vs Right Hand	51.1%
Subject 20	MI Left vs Right Hand	51.1%
Subject 21	MI Left vs Right Hand	53.3%
Subject 22	MI Left vs Right Hand	57.7%
Subject 23	MI Left vs Right Hand	48.8%
Subject 24	MI Left vs Right Hand	53.3%
Subject 25	MI Left vs Right Hand	48.8%
Subject 26	MI Left vs Right Hand	71.1%
Subject 27	MI Left vs Right Hand	51.1%
Subject 28	MI Left vs Right Hand	48.8%
Subject 29	MI Left vs Right Hand	93.3%
Subject 30	MI Left vs Right Hand	60%
Subject 31	MI Left vs Right Hand	80%
Subject 32	MI Left vs Right Hand	73.3%
Subject 33	MI Left vs Right Hand	57.7%
Subject 34	MI Left vs Right Hand	71.1%
Subject 35	MI Left vs Right Hand	75.5%
Subject 36	MI Left vs Right Hand	75.5%
Subject 37	MI Left vs Right Hand	51.1%
Subject 38	MI Left vs Right Hand	48.8%
Subject 39	MI Left vs Right Hand	48.8%
Subject 40	MI Left vs Right Hand	73.3%
Subject 41	MI Left vs Right Hand	55.5%
Subject 42	MI Left vs Right Hand	82.2%
Subject 43	MI Left vs Right Hand	60%
Subject 44	MI Left vs Right Hand	62.2%
Subject 45	MI Left vs Right Hand	64.4%
Subject 46	MI Left vs Right Hand	55.5%
Subject 47	MI Left vs Right Hand	48.8%
Subject 48	MI Left vs Right Hand	64.4%
Subject 49	MI Left vs Right Hand	62.2%
Subject 50	MI Left vs Right Hand	53.3%
Subject 51	MI Left vs Right Hand	48.8%
Subject 52	MI Left vs Right Hand	66.6%
Subject 53	MI Left vs Right Hand	84.4%
Subject 54	MI Left vs Right Hand	64.4%
Subject 55	MI Left vs Right Hand	64.4%
Subject 56	MI Left vs Right Hand	95.5%
Subject 57	MI Left vs Right Hand	71.1%
Subject 58	MI Left vs Right Hand	66.6%
Subject 59	MI Left vs Right Hand	60%
Subject 60	MI Left vs Right Hand	66.6%
Subject 61	MI Left vs Right Hand	64.4%
Subject 62	MI Left vs Right Hand	75.5%
Subject 63	MI Left vs Right Hand	60%
Subject 64	MI Left vs Right Hand	48.8%
Subject 65	MI Left vs Right Hand	57.7%
Subject 66	MI Left vs Right Hand	48.8%
Subject 67	MI Left vs Right Hand	53.3%
Subject 68	MI Left vs Right Hand	51.1%
Subject 69	MI Left vs Right Hand	88.8%
Subject 70	MI Left vs Right Hand	64.4%
Subject 71	MI Left vs Right Hand	48.8%
Subject 72	MI Left vs Right Hand	86.6%
Subject 73	MI Left vs Right Hand	48.8%
Subject 74	MI Left vs Right Hand	48.8%
Subject 75	MI Left vs Right Hand	51.1%
Subject 76	MI Left vs Right Hand	48.8%
Subject 77	MI Left vs Right Hand	55.5%
Subject 78	MI Left vs Right Hand	53.3%
Subject 79	MI Left vs Right Hand	48.8%
Subject 80	MI Left vs Right Hand	48.8%
Subject 81	MI Left vs Right Hand	64.4%
Subject 82	MI Left vs Right Hand	66.6%
Subject 83	MI Left vs Right Hand	71.1%
Subject 84	MI Left vs Right Hand	57.7%
Subject 85	MI Left vs Right Hand	86.6%
Subject 86	MI Left vs Right Hand	51.1%
Subject 87	MI Left vs Right Hand	48.8%
Subject 90	MI Left vs Right Hand	48.8%
Subject 91	MI Left vs Right Hand	53.3%
Subject 93	MI Left vs Right Hand	48.8%
Subject 94	MI Left vs Right Hand	84.4%
Subject 95	MI Left vs Right Hand	44.4%
Subject 96	MI Left vs Right Hand	66.6%
Subject 97	MI Left vs Right Hand	62.2%
Subject 98	MI Left vs Right Hand	48.8%
Subject 99	MI Left vs Right Hand	55.5%
Subject 101	MI Left vs Right Hand	84.4%
Subject 102	MI Left vs Right Hand	62.2%
Subject 103	MI Left vs Right Hand	66.6%
Subject 105	MI Left vs Right Hand	62.2%
Subject 107	MI Left vs Right Hand	57.7%
Subject 108	MI Left vs Right Hand	48.8%
Subject 109	MI Left vs Right Hand	48.8%

Table A.1: Leave-One-Out inter-subject (personalised) performance in the MI-classification (Right / Left Hand) task in PhysioNet dataset for BrainWave-Scattering Net (ours).

Subject	Task	Accuracy	Subject	Task	Accuracy
Subject 1	MI Left vs Rest	64.9%	Subject 67	MI Left vs Rest	88.7%
Subject 2	MI Left vs Rest	64.7%	Subject 68	MI Left vs Rest	90.3%
Subject 3	MI Left vs Rest	50.3%	Subject 69	MI Left vs Rest	50.3%
Subject 4	MI Left vs Rest	50.3%	Subject 70	MI Left vs Rest	62.9%
Subject 5	MI Left vs Rest	52.0%	Subject 71	MI Left vs Rest	70.3%
Subject 6	MI Left vs Rest	57.3%	Subject 72	MI Left vs Rest	87.3%
Subject 7	MI Left vs Rest	82.9%	Subject 73	MI Left vs Rest	64.7%
Subject 8	MI Left vs Rest	68.5%	Subject 74	MI Left vs Rest	55.6%
Subject 9	MI Left vs Rest	77.8%	Subject 75	MI Left vs Rest	51.8%
Subject 10	MI Left vs Rest	64.5%	Subject 76	MI Left vs Rest	64.9%
Subject 11	MI Left vs Rest	73.8%	Subject 77	MI Left vs Rest	53.8%
Subject 12	MI Left vs Rest	68.7%	Subject 78	MI Left vs Rest	59.6%
Subject 13	MI Left vs Rest	52.3%	Subject 79	MI Left vs Rest	64.7%
Subject 14	MI Left vs Rest	92.7%	Subject 80	MI Left vs Rest	68.5%
Subject 15	MI Left vs Rest	59.2%	Subject 81	MI Left vs Rest	72.0%
Subject 16	MI Left vs Rest	57.4%	Subject 82	MI Left vs Rest	55.2%
Subject 17	MI Left vs Rest	60.9%	Subject 83	MI Left vs Rest	57.1%
Subject 18	MI Left vs Rest	59.3%	Subject 84	MI Left vs Rest	59.4%
Subject 19	MI Left vs Rest	50.3%	Subject 85	MI Left vs Rest	66.9%
Subject 20	MI Left vs Rest	77.8%	Subject 86	MI Left vs Rest	50.3%
Subject 21	MI Left vs Rest	65.4%	Subject 87	MI Left vs Rest	72.0%
Subject 22	MI Left vs Rest	74.2%	Subject 88	MI Left vs Rest	61.3%
Subject 23	MI Left vs Rest	77.8	Subject 89	MI Left vs Rest	63.1%
Subject 24	MI Left vs Rest	81.4%	Subject 90	MI Left vs Rest	50.3%
Subject 25	MI Left vs Rest	62.7%	Subject 91	MI Left vs Rest	61.4%
Subject 26	MI Left vs Rest	76.2%	Subject 92	MI Left vs Rest	53.6%
Subject 27	MI Left vs Rest	75.8%	Subject 93	MI Left vs Rest	57.3%
Subject 28	MI Left vs Rest	64.7%	Subject 94	MI Left vs Rest	68.2%
Subject 29	MI Left vs Rest	50.3%	Subject 95	MI Left vs Rest	57.4%
Subject 30	MI Left vs Rest	68.2%	Subject 96	MI Left vs Rest	63.1%
Subject 31	MI Left vs Rest	55.3%	Subject 97	MI Left vs Rest	50.3%
Subject 32	MI Left vs Rest	71.8%	Subject 98	MI Left vs Rest	92.7%
Subject 33	MI Left vs Rest	59.3%	Subject 99	MI Left vs Rest	74.3%
Subject 34	MI Left vs Rest	81.8%	Subject 100	MI Left vs Rest	79.3%
Subject 35	MI Left vs Rest	63.1%	Subject 101	MI Left vs Rest	63.1%
Subject 36	MI Left vs Rest	65.1%	Subject 102	MI Left vs Rest	74.0%
Subject 37	MI Left vs Rest	64.7%	Subject 103	MI Left vs Rest	59.4%
Subject 38	MI Left vs Rest	90.7%	Subject 104	MI Left vs Rest	87.1%
Subject 39	MI Left vs Rest	81.4%	Subject 105	MI Left vs Rest	62.3%
Subject 40	MI Left vs Rest	68.2%	Subject 106	MI Left vs Rest	68.3%
Subject 41	MI Left vs Rest	59.1%	Subject 107	MI Left vs Rest	76.0%
Subject 42	MI Left vs Rest	55.3%	Subject 108	MI Left vs Rest	52.0%
Subject 43	MI Left vs Rest	70.3%	Subject 109	MI Left vs Rest	59.1%
Subject 44	MI Left vs Rest	50.3%	Subject 110	MI Left vs Rest	55.3%
Subject 45	MI Left vs Rest	50.3%	Subject 111	MI Left vs Rest	64.5%
Subject 46	MI Left vs Rest	50.3%	Subject 112	MI Left vs Rest	50.3%
Subject 47	MI Left vs Rest	81.4%	Subject 113	MI Left vs Rest	50.3%
Subject 48	MI Left vs Rest	52.2%	Subject 114	MI Left vs Rest	55.8%
Subject 49	MI Left vs Rest	60.7%	Subject 115	MI Left vs Rest	77.8%
Subject 50	MI Left vs Rest	90.7%	Subject 116	MI Left vs Rest	83.3%
Subject 51	MI Left vs Rest	83.6%	Subject 117	MI Left vs Rest	81.8%
Subject 52	MI Left vs Rest	70.3%	Subject 118	MI Left vs Rest	57.3%

Table A.2: 5-fold cross-validation inter-subject (personalised) performance in the MI-classification (Left Hand / Rest) task in MyBrainCommands - Lab Setting dataset for BrainWave-Scattering Net for 104 random subjects (ours).