Image-Based Motor Imagery EEG Classification With Deep Learning for Brain Computer Interfaces

Ward Fadel

Pázmány Péter Catholic University

Faculty of Information Technology and Bionics

Roska Tamás Doctoral School of Sciences and Technology

Supervisors

Dr. István Ulbert, DSc

Dr. Lucia Wittner, DSc

A thesis submitted for the degree of Doctor of Philosophy

Budapest, 2025

To the Beauty of Life

Acknowledgments

To Hungary, and the environment where I was able to get to know myself more deeply and discover what the word *"other"* truly means through its multispaces.

To Prof. István Ulbert, the exceptional spirit who created for us the atmosphere to grow and discover — I will always be thankful for your generous free mind, patience, and wonderful humanity. To Dr. Lucia Wittner, who supported me greatly throughout my scientific journey.

To my colleagues, collaborators, and administrative members at UlbertLab and the Institute of Cognitive Neuroscience and Psychology, HUN-REN Research Centre for Natural Sciences, I am grateful for the honest discussions and the heartfelt moments we shared. I wish to thank the Cybathlon pilots and Brain-Computer Interfacing team for the invaluable opportunity to be part of such an inspiring experience, as well as Dr. Csaba Köllőd, András Adolf, Dr. Ágnes Kandrács, Dr. Bálint File, Dr. Csaba Horváth, Dr. Domokos Meszéna, Dr. Domonkos Horváth, Eszter Juhász, Dr. Estilla Zsófia Tóth, Dr. Gergely Márton, Dr. János Rokai, Dr. Kinga Tóth, Klára Malikné (for the very kind help), Levente Víg, Dr. Melinda Rácz, Mihály Rámpay, Dr. Moutz Wahdow, Nour Essam, Péter Seré, Rebeka Stelcz, Réka Bod, Dr. Richárd Fiáth, Dr. Tibor Nánási , and Dr. Zoltán Somogyvári.

To the Roska Tamás Doctoral School of Sciences and Technology, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University (ITK-PPKE), for all the support throughout these past years. I am grateful to the professors and administrative members: Professor Árpád Csurgay, for guiding us through the fascinating journey of quantum worlds and for your kind help and encouragement, Dr. Vida Tivadarné Katinka, for her immense support and care, Dr. Helga Reguly, for mentoring the early startup projects, Dr. András Horváth, Dr. Iván Kristóf, Prof. György Cserey, Prof. Géza Kolumbán, Prof. Gábor Szederkényi, and Prof. Péter Szolgay, and the staff of the Registrar's Office, Dean's Office and the Financial Department.
To my colleagues and friends at PPKE, Dr. Sam Khozama, Dr. Yahya Ibrahim, Dr. Jalal Alafandi, Dr. Nawar Alhemeary, Dr. Janos Juhasz, Dr. Mate Mohacsi, Dr. Ágnes Szabó, and Dr. Suchana Chakravarty, and to many others I am greatful for the wonderful times we spent together.
To my childhood friends who are like brothers and sisters, to my neighbors, to my teachers, to my colleagues, and to 'Fairouz voice' whose presence has shaped me in countless ways.

To my family, to the powerful and beloved soul of my grandmother Sadika, who believed in excellence and infinite love and respect, to my grandfather Fadel, whose presence continues to guide me, to my beloved aunts and uncles who cared for me since childhood, to my beloved cousins, to my beloved brothers, to my mother Jenan, the sweetest heart on earth who keeps us in her prayers every day, and to my father Ghassan, the exceptional person whose loving mazes continues to inspire me, and to the 'Laughing Dove' over his window.

To my country, Syria, and the eternal sea and the poets — those light shepherds, those distance shepherds — of my hometown Jablah, I dedicate this humble early work — with the belief that every delay may carry within it a hidden goodness.

My research drew upon studies I had published during the early stages of my PhD, while more recent work was postponed due to the dark period I went through in 2024, a period that may shed light on some darkness and open new worlds for me. The focus was on transforming motor-imagery EEG signals into consecutive images that preserve neurophysiological meaning (video-like) and applying deep learning for classification. This approach remains valid and continues to gain attention. Additionally, I explored a non-canonical frequency band (Delta), which showed improved results. Although further validation is needed, recent work by my colleague András Adolf supports the significance of the Delta band for motor imagery classification.

Finally, I hope that the classification of brain states will serve as a true threshold to explore the unthinkable, with minimal side effects.

Table of Contents

Contents

List of Figures

Abbreviations

BCI - Brain Computer Interface

BPTT - Backpropagation Through Time

CNN - Convolutional Neural Network

CNV - Contingent Negative Variation

CMRR - Common Mode Rejection Ratio

CSP - Common Spatial Patterns

DL - Deep Learning

DFFT - Discrete Fast Fourier Transform

ECG – Electrocardiography

EEG – Electroencephalography

EMG – Electromyography

EP - Evoked Potential

ERP - Event-related Potential

ERD - Event Related Desynchronization

ERS - Event Related Synchronization

EPSPs - Excitatory Postsynaptic Potentials

FFT - Fast Fourier Transform

IPSPs - Inhibitory Postsynaptic Potentials

KMI - Kinesthetic Motor Imagery

KMV - Visual Motor Imagery

LSTM - Long-Short Term Memory

MI - Motor Imagery

ML - Machine Learning

MRI - Magnetic Resonance Imaging

NN - Neural Network

NIRS - Near-Infrared Spectroscopy

PSPs - Postsynaptic Potentials

PMBR – Post Movement Beta Rebound

PSD - Power Spectral Density

RNN - Recurrent Neural Network

SCP - Slow Cortical Potentials

SNR - Signal to Noise Ratio

STFT - Short-Time Fourier Transform

VEP - Visually Evoked Potential

WT - Wavelet Transform

SSVEP - Steady-State Visually Evoked Potential

Chapter 1

Introduction

1.1 Motivation and Goals

Electroencephalography (EEG)-based Motor Imagery (MI) classification forms the backbone of non-invasive Brain–Computer Interface (BCI) systems, offering users the ability to interact with external devices solely through imagined movements. This has profound implications for assistive technologies, rehabilitation, and human–computer interaction. Despite decades of research, however, robust, generalizable, and high-performance MI-EEG classifiers remain elusive. The difficulty arises from the intrinsic characteristics of EEG: its low signal-to-noise ratio, coarse spatial resolution, and strong inter- and intra-subject variability. These limitations make reliable decoding of motor imagery highly challenging in practical, real-world settings.

Traditional BCI pipelines have largely relied on handcrafted features—such as bandpower modulations in Mu (8–13 Hz) and Beta (13–30 Hz) rhythms at electrodes C3, C4, and Cz, typically associated with event-related desynchronization (ERD) and synchronization (ERS). While such features capture the canonical sensorimotor rhythm, they fail to consistently generalize. Inter- and intra-session variability, deviations from expected ERD/ERS patterns, and the inherently dynamic and distributed nature of cortical activations mean that C3/C4/Cz-centric approaches are insufficient. Moreover, these pipelines frequently employ shallow classifiers such as Linear Discriminant Analysis (LDA) or Support Vector Machines (SVM), which cannot fully exploit the complex spatiotemporal structure of EEG.

Recent advances in deep learning offer a powerful alternative: rather than manually crafting features, neural networks can learn directly from data, provided the input representation is rich and structured. However, EEG signals are not naturally image-like or spatially uniform, which poses a barrier to leveraging state-of-the-art deep learning architectures from computer vision. This motivates a rethinking of how EEG should be represented before classification.

The central motivation of this thesis is therefore twofold:

1- To overcome the limitations of traditional electrode- and feature-centric methods by moving beyond sparse sensorimotor channels and embracing the full 64-electrode montage, thereby capturing richer spatial patterns across the cortex.
2- To enable deep learning to operate effectively on EEG data by transforming irregular, noisy signals into structured 2D image representations that preserve neurophysiological meaning while aligning with convolutional neural network (CNN) inductive biases.

Guided by these motivations, this work pursues the following goals:

1- Design novel EEG-to-image transformations. We propose two complementary methods:
   - Chessboard Projection, introduced here as a novel discrete grid-aligned encoding, which maps electrodes onto a regularized 2D lattice optimized for compatibility with CNNs.

- Interpolated Azimuthal Projection, deployed here for MI-EEG, which preserves cortical topology through geodesic projection and interpolation, producing smooth topographic maps.

2- Develop a hybrid spatiotemporal deep learning framework. A CNN branch is designed to extract spatial and spectral features from the transformed images, while an LSTM branch captures the temporal dynamics of cortical activity across motor imagery epochs. This architecture is co-designed with the input representation, scaled to exploit the distinct properties of Chessboard and Azimuthal encodings.
3- Systematically evaluate the role of spectral bands. Beyond Mu and Beta, we investigate the contribution of the Delta band (0.5–4 Hz). Our analysis challenges the prevailing assumption of Delta's irrelevance in motor imagery, demonstrating that its decodability is contingent on spatial encoding.
4- Validate on a large-scale dataset. We employ a 64-channel, multi-subject dataset (103 out of 109 subjects utilized) across multiple MI classes, ensuring that findings are not limited to small, constrained experimental setups. This allows us to assess both within-subject and cross-subject generalizability, a critical step toward practical BCI deployment.
5- Reframe representation as a decisive factor in EEG decoding. By contrasting Chessboard and Azimuthal projections, we show that spatial encoding is not a neutral preprocessing step but a design choice that governs what information the brain makes accessible to machine learning.

In summary, the motivation driving this thesis is the inadequacy of conventional MI-EEG approaches in handling variability, complexity, and distributed cortical dynamics. The goal is to establish image transformation as a principled and powerful paradigm for EEG representation, enabling deep learning models to achieve more robust, interpretable, and generalizable motor imagery classification.

1.2 Thesis Overview

This thesis is organized into five chapters, progressing from theoretical foundations to methodological development, empirical evaluation, and final synthesis. The structure is designed to guide the reader from the neurophysiological basis of EEG signals to the introduction of novel image-based representations and deep learning frameworks for Motor Imagery (MI) classification.

Chapter 2: Background and Theory

This chapter provides the neurophysiological and methodological foundations necessary to understand EEG and its role in Brain–Computer Interfaces (BCIs). It begins with a historical overview of EEG and the neuronal mechanisms underlying electrophysiological recordings, including the generation of postsynaptic potentials, volume conduction, and oscillatory dynamics. The discussion then contrasts invasive and non-invasive neuroimaging methods, situating EEG within the broader context of brain measurement techniques.

Chapter 3: Brain–Computer Interfacing Using Motor Imagery EEG

Here, the thesis narrows its focus to BCI systems, with an emphasis on Motor Imagery as a widely studied paradigm. After surveying different BCI approaches—such as P300, SSVEP, SCP, and error-related potentials—the chapter delves into MI-based BCIs, discussing their theoretical underpinnings, challenges, and relevance for non-invasive neural control.

Chapter 4: Signal-to-Image Transformation and Deep Learning
This chapter constitutes the methodological core of the thesis. It first introduces deep learning architectures relevant to EEG decoding, including Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), and Long Short-Term Memory (LSTM) networks. Subsequently, it presents two complementary EEG-to-image transformation methods: Interpolated Azimuthal Projection, which preserves cortical topology through geodesic mapping and interpolation, and Chessboard Projection, introduced as a novel discrete electrode-to-grid mapping optimized for CNN processing. The chapter further details the experimental setup, including the use of the PhysioNet Motor Movement/Imagery dataset, preprocessing procedures, and the hybrid CNN–LSTM architecture designed for spatiotemporal learning.

Chapter 5: Summary and Scientific Results
The final chapter synthesizes the findings into two major contributions:

Thesis Group I introduces and evaluates the two signal-to-image transformation pipelines.

Thesis Group II demonstrates the effectiveness of an end-to-end hybrid CNN–LSTM framework for spatiotemporal MI decoding, including the novel finding that Delta-band dynamics contribute to classification accuracy when spatially preserved.
The chapter concludes with a discussion of publications arising from this work and perspectives for future research directions.

Chapter 2

Background and Theory

2.1 Historical Foundations of EEG

The origins of electroencephalography (EEG) can be traced to the 19th century with the discovery of bioelectricity. In 1848, Emil du Bois-Reymond demonstrated that nerves and muscles produce electrical discharges, establishing electrophysiology as a scientific discipline and proving that biological tissues are electrically as well as chemically active [1]. Building on this foundation, Richard Caton in 1875 recorded spontaneous fluctuations of cerebral potentials in rabbits and monkeys using galvanometers. He noted that these signals varied with sensory stimulation and sleep, foreshadowing the later connection between brain states and measurable electrical activity [1,2]. Subsequent work expanded these findings. In the 1880s, Adolf Beck performed cortical recordings in dogs and rabbits, observing rhythmic oscillations that diminished with sensory stimulation, an early description of what would later be termed desynchronization [4]. In 1912, Vladimir Pravdich-Neminsky published the first mammalian electroencephalogram, showing both spontaneous rhythms and evoked responses [5]. Despite technological limitations and the disruptions of World War I, these pioneering studies created the scientific lineage that led to the first human EEG.

The decisive breakthrough came with Hans Berger, who in 1924 recorded rhythmic oscillations from the human scalp using a string galvanometer and silver wire electrodes. He coined the term *Elektrenkephalogramm* and in 1929 described two distinct patterns: the alpha rhythm (~10 Hz), prominent in the occipital regions during relaxed wakefulness, and faster Beta activity linked to alertness [6]. Berger's work initially faced skepticism, as many suspected the signals were artifacts of muscle or cardiac activity [1]. However, independent confirmation by Adrian and Matthews in 1934 validated his findings and demonstrated the suppression of alpha rhythm by eye opening [7]. By the late 1930s, EEG had spread internationally and shown clear clinical value, particularly in epilepsy, where Berger himself had documented abnormal discharges [8]. Today Berger is universally recognized as the founder of human EEG.

Following Berger's work, EEG rapidly expanded during the 1930s and 1940s. Gibbs, Davis, and Lennox in the U.S. demonstrated the characteristic spike-and-wave discharges of absence epilepsy, establishing EEG as the gold standard for seizure diagnosis [9]. In parallel, EEG studies of sleep identified distinct stages and revealed the existence of rapid eye movement (REM) sleep, laying the foundation for modern sleep medicine [10].

Technological advances—such as vacuum-tube amplifiers, oscilloscopes, and continuous paper recording—greatly improved signal sensitivity and reliability [11]. At the same time, Herbert Jasper introduced the 10–20 system of electrode placement in 1958, providing a standardized framework for electrode positioning that remains in use today [12]. By the mid-20th century, EEG was firmly established as a core method in clinical neurophysiology, applied not only to epilepsy and sleep but also to head trauma, encephalopathies, degenerative diseases, and brain death assessment [1]. These developments ensured EEG's enduring role as both a diagnostic tool and a window into human brain function.

2.2 Neuronal Basis of EEG

The electroencephalogram ultimately reflects the collective activity of neurons, particularly in the cerebral cortex. Understanding how microscopic cellular processes give rise to macroscopic scalp signals requires an examination of neuronal structure, the basic principles of excitability, and the integration of synaptic activity across populations of neurons.

2.2.1 Neuronal Structure and Excitability

Neurons are the fundamental units of the nervous system. Each neuron is composed of a cell body (soma), dendrites, and an axon. The soma contains the nucleus and metabolic machinery, while dendrites serve as the primary input structures, receiving signals from thousands of presynaptic neurons. The axon is the primary output pathway, capable of transmitting signals over long distances to other neurons or target tissues [13]. Electrical excitability in neurons arises from the distribution of ions across the cell membrane. At rest, the neuronal membrane maintains a negative potential relative to the extracellular space due to the asymmetric distribution of sodium, potassium, chloride, and calcium ions. This resting membrane potential is sustained by selective ion permeability and the activity of transport mechanisms such as the sodium–potassium pump. Small changes in ion flux across the membrane are sufficient to alter excitability and drive communication between neurons [14]. To fully grasp the staggering complexity of these microcircuits at the single-cell level, recent advances in nanoscale connectomics now make it possible to visualize — for the first time in human tissue — the precise structural convergence of inputs onto individual neurons. Single-neuron reconstructions from human cortical tissue reveal how thousands of axons converge onto individual cells, forming dense networks of excitatory and inhibitory inputs. Figure 1 illustrates both the incoming axons and the neuron's own synapses, offering insight into the structural basis of neural computation.

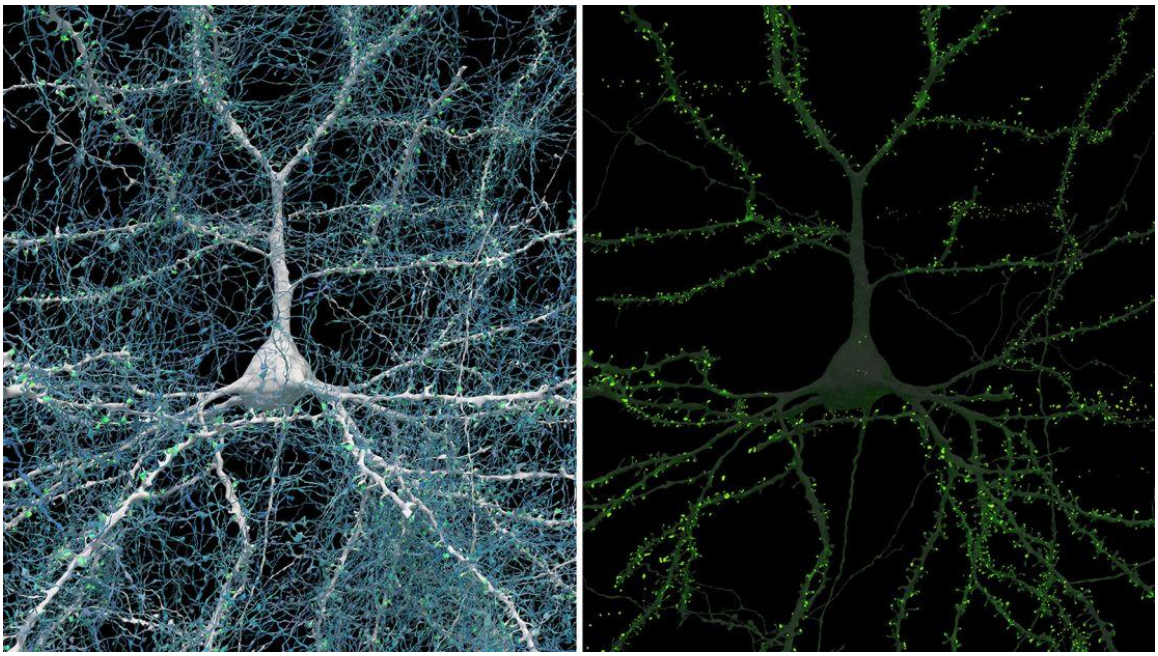

Figure 1: Left: 3D reconstruction of a single human neuron (white; soma ~14 μm in diameter) receiving ~5,600 incoming axons, shown in blue, with synaptic contacts highlighted in green. Right: the same neuron's synapses displayed in green. The human brain contains an estimated 86 billion neurons. (Google Research & Harvard University [15])

2.2.2 Action Potentials and Postsynaptic Potentials

When a neuron is sufficiently depolarized, it generates an action potential, a rapid all-or-none electrical event that propagates along the axon. Action potentials are crucial for long-distance signaling within the nervous system and represent one source of bioelectrical activity. However, they contribute only minimally to scalp EEG because they are too brief and localized to summate effectively at the level of populations [1]. In contrast, postsynaptic potentials (PSPs) are the primary contributors to EEG signals. These potentials arise when neurotransmitters bind to receptors on dendrites or the soma, leading to ion channel opening and subsequent changes in membrane potential. Excitatory postsynaptic potentials (EPSPs) increase the likelihood of neuronal firing, typically through depolarization, while inhibitory postsynaptic potentials (IPSPs) reduce excitability, often via hyperpolarization. Unlike action potentials, PSPs are graded in amplitude, last longer in duration, and occur over broad dendritic fields. This makes them more suitable for spatial and temporal summation across large ensembles of neurons, which is essential for generating the extracellular fields measurable on the scalp [15,16].

2.2.3 Synaptic Integration and Population Activity

EEG does not reflect the activity of individual neurons, but rather the coordinated activity of large neuronal populations. Cortical pyramidal neurons, in particular, play a dominant role due to their abundance, geometry, and orientation. Their apical dendrites extend perpendicularly to the cortical surface, aligning them in parallel arrays. When these neurons undergo synchronous synaptic activity, the resulting currents sum to form macroscopic dipoles that are detectable outside the skull [1]. Integration occurs both within individual neurons and across neural populations. Within a single neuron, excitatory and inhibitory inputs are integrated at the soma, influencing the probability of action potential initiation. At the population level, columns of pyramidal neurons within cortical layers act as functional units. Figure. 2 illustrates the general connectivity of a neocortical column, as originally drawn by the pioneering neuroanatomist János Szentágothai. The diagram depicts the canonical six-layered structure of the neocortex, labeled with Roman numerals I through VI, each layer housing distinct neuronal populations and connection patterns. Pyramidal neurons — the principal excitatory cells of the cortex — span multiple layers and form the backbone of intra- and inter-columnar communication. Local circuits within the column, along with long-range horizontal and vertical projections, enable synchronization across neuronal ensembles. This synchronized activity underlies the rhythmic fluctuations in membrane potentials observed at the population level, which are fundamental to cortical computation, information integration, and the generation of brain rhythms such as alpha, Beta, and gamma oscillations [17]. In summary, while the action potential is fundamental for neural communication, it is primarily the summated postsynaptic potentials of aligned cortical pyramidal neurons that dominate the scalp EEG signal. The geometry of these neurons, their columnar organization, and their ability to synchronize across networks provide the structural and physiological substrate for the generation of measurable brain rhythms. Szentágothai's schematic remains a foundational reference for understanding cortical microcircuitry and its role in large-scale network dynamics [18].

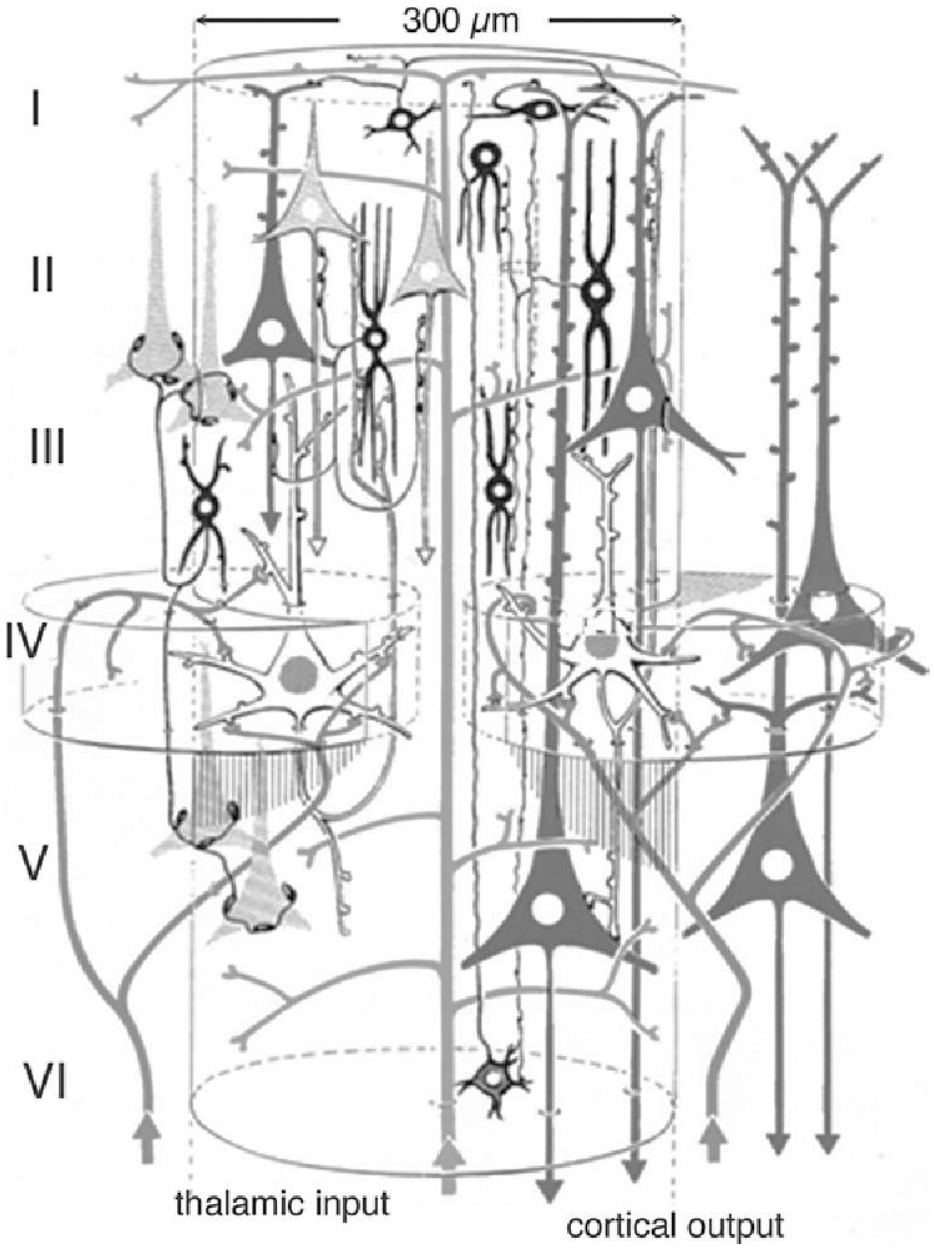

Figure 2: General Connectivity of a neocortical column drawn by Janos Szentágothai. Roman numerals (I to VI) refer to the six cortical layers [18].

To visualize the anatomical substrate underlying these population-level dynamics, recent advances in connectomics now offer unprecedented resolution. Figure. 3 presents a high-resolution 3D reconstruction of a ~1 mm^3 human cortical sample, revealing the laminar distribution of excitatory neurons — generated through the Harvard–Google connectomics collaboration. Color-coded by depth from the cortical surface (blue = superficial, fuchsia = deep), this reconstruction maps approximately 57,000 excitatory neurons, 230 millimeters of vasculature, and 150 million synapses. Released in May 2024 and representing the most detailed human cortical reconstruction to date (as of April 2025), this dataset provides a nanoscale anatomical foundation for understanding how laminar organization and neuronal alignment support synchronized activity detectable by EEG.

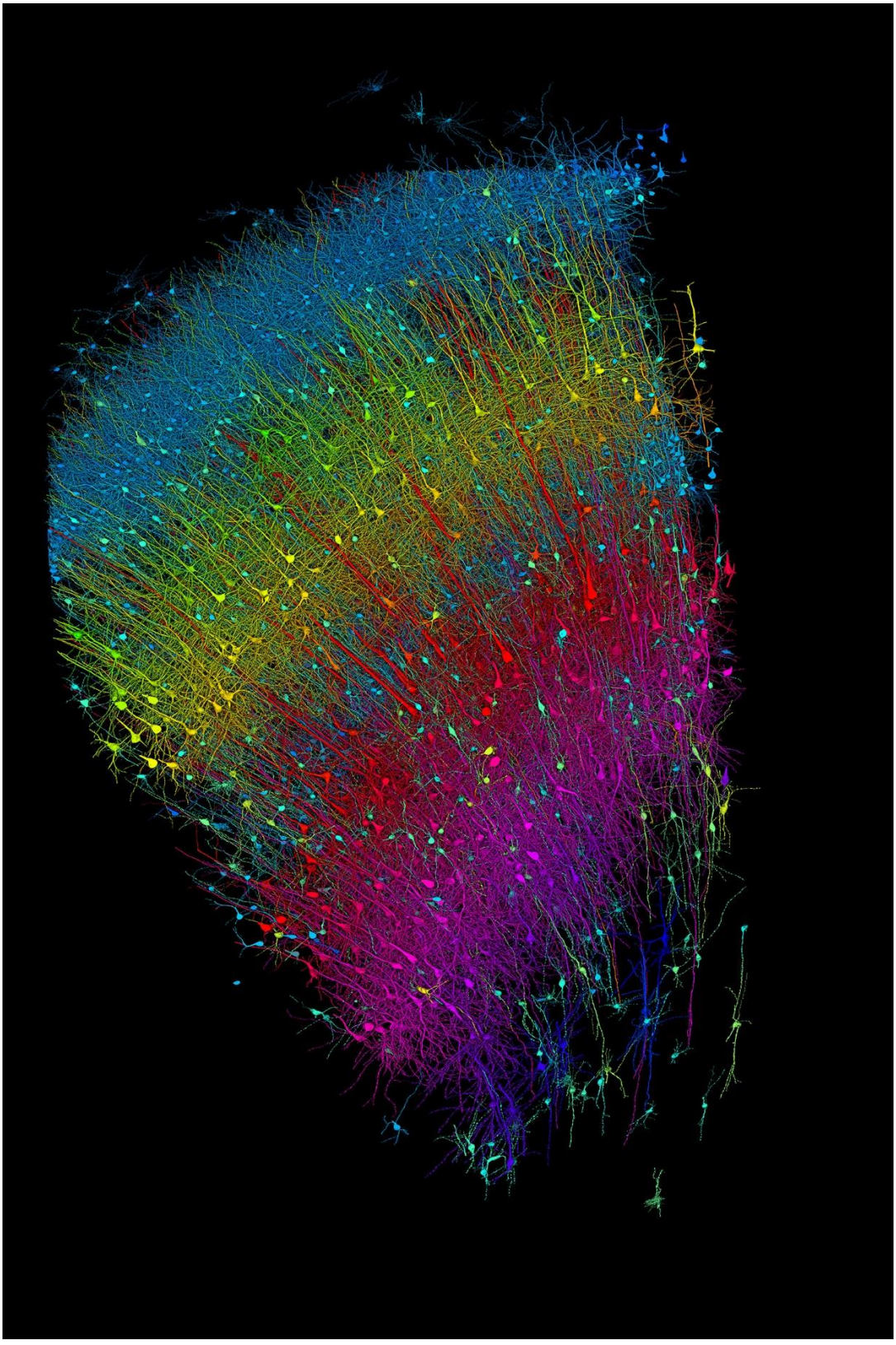

Figure 3: Excitatory neurons color-coded by depth from the cortical surface, from blue (superficial) to fuchsia (deep), within a ~1 mm³ tissue sample (≈57,000 cells, ~230 millimeters of blood vessels and ~150 million synapses). This represents the most detailed human brain reconstruction to date (August 2025), released in May 2024 by Google Research & Harvard University [15].

2.2.4 Functional Organization of the Sensorimotor Cortex

The synchronized activity of neuronal populations, described in the previous subsection, provides the foundation for understanding how the cortex encodes sensory and motor functions at a macroscopic level. Large populations of pyramidal neurons, organized into cortical columns and aligned across layers, generate the population-level signals measurable by EEG. These neuronal populations are also topographically organized across distinct cortical regions, often delineated by Brodmann areas, which correspond to primary sensory, motor, and associative functions (Figure 4). For example, areas 1–3 correspond to the primary somatosensory cortex, area 4 to the primary motor cortex, area 17 to the primary visual cortex, and areas 44/45 to Broca's language region.

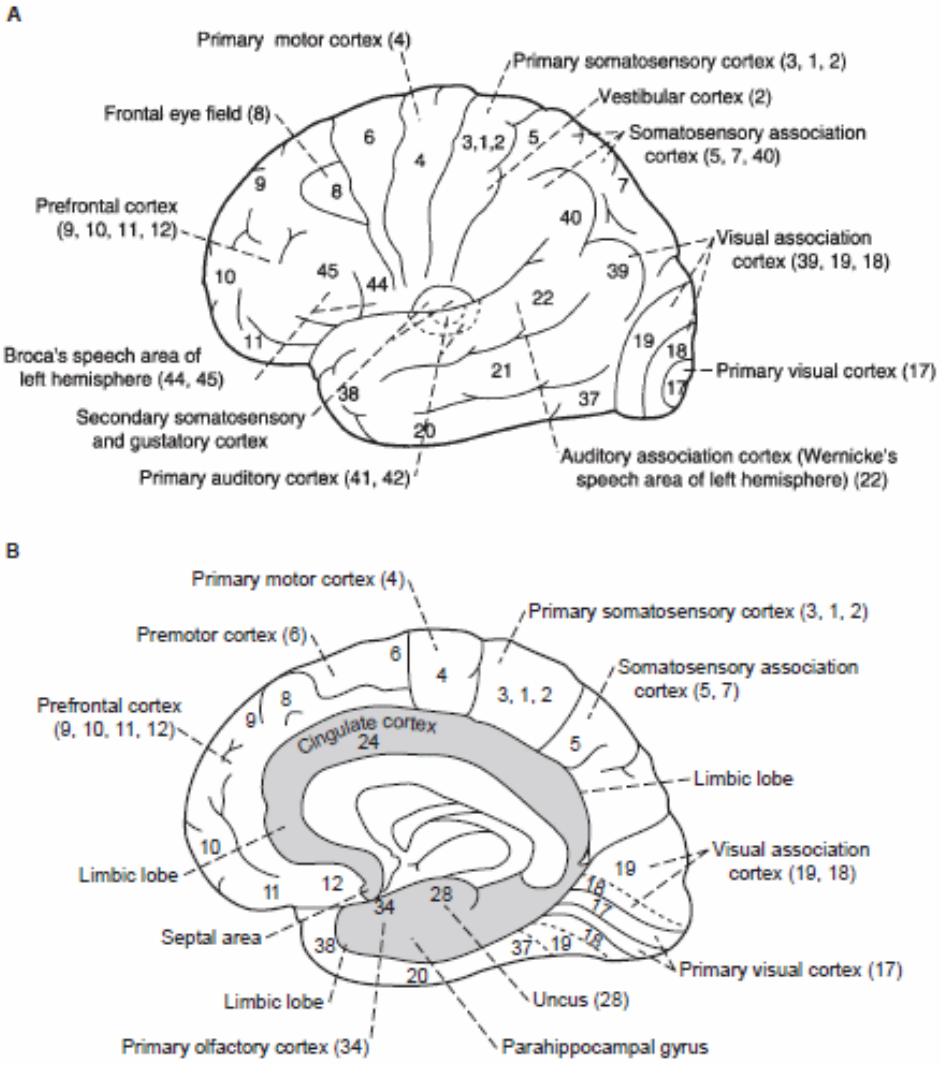

Figure 4: Brodmann Areas of the Human Cortex. Schematic map of the cerebral cortex showing selected Brodmann areas and their approximate functional assignments, with (a) lateral and (b) medial views of the cortical surface and inner structures [19].

Building on this anatomical framework, Penfield's cortical homunculi provide a functional illustration of how body regions are represented across the sensorimotor cortex. The somatosensory homunculus (A) on the postcentral gyrus highlights regions with high sensory acuity, such as the lips, tongue, and fingertips, whereas the motor homunculus (B) on the

precentral gyrus emphasizes areas requiring fine motor control, notably the hands and face (Figure 5). These representations reflect the principle that cortical allocation is determined by functional importance rather than physical size. Recent neuroimaging and electrophysiological studies indicate that cortical boundaries and proportional representations are more complex and variable across individuals, with overlapping somatotopic patterns.

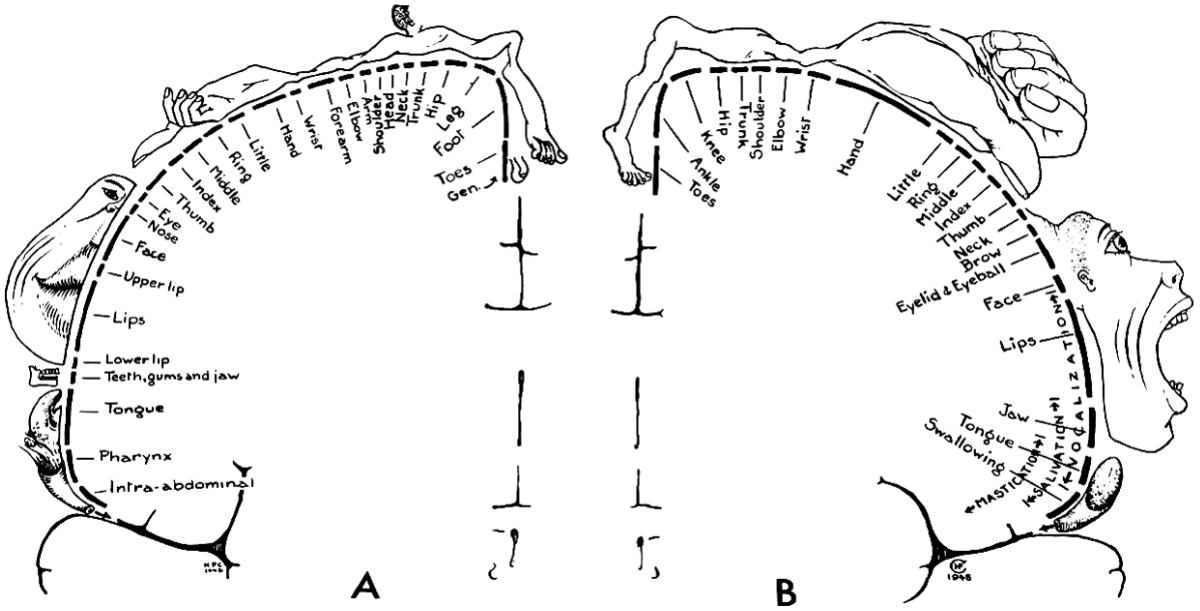

Figure 5: The cortical homunculi originally mapped by Wilder Penfield in 1948. A: the somatosensory homunculus on the postcentral gyrus of the parietal lobe, illustrating cortical representation of touch and sensation. B: the motor homunculus on the precentral gyrus of the frontal lobe, illustrating cortical representation of voluntary movement. Disproportionately large hands, lips, and tongue reflect their greater cortical allocation.

2.3 From Local Currents to Scalp Potentials

2.3.1 Current Sources and Dipole Formation

The fundamental units of EEG signals are current sources and sinks created by synaptic activity. When neurotransmitters open ion channels on the dendritic membrane of pyramidal neurons, localized inward and outward currents are established. These currents generate extracellular fields that can be approximated as dipoles, consisting of a source (outward current) and a sink (inward current) [14,15].

Cortical pyramidal neurons are particularly suited to generating measurable dipoles because of their geometry and orientation. Their apical dendrites extend perpendicularly toward the cortical surface, and they are arranged in parallel within cortical columns. This alignment creates an "open-field" configuration, in which individual dipoles reinforce one another rather than canceling out. In contrast, neurons with more spherical or random orientations form "closed-field" configurations that produce little or no contribution to scalp EEG [15].

The synchronous activation of large populations of pyramidal neurons thus produces macroscopic dipole layers. These dipoles sum across cortical patches spanning several square centimeters to generate extracellular fields strong enough to be detected by scalp electrodes.

2.3.2 Volume Conduction in Brain Tissue

Once generated, these extracellular fields propagate through the brain and surrounding tissues via volume conduction. The cerebral cortex, cerebrospinal fluid (CSF), skull, and scalp each act as conductive media, shaping the spatial profile of the potentials that ultimately reach the electrodes [14].

Because brain tissue and CSF are relatively good conductors, currents spread broadly before encountering the skull, which has high resistance and low conductivity. The skull attenuates signal amplitude by an order of magnitude or more and spatially smears the fields, reducing the ability to resolve fine-scale sources [11]. The scalp, being more conductive, provides less additional distortion, but the combined effect of these layers is that EEG signals recorded at the surface represent a blurred and attenuated version of the underlying cortical activity.

This property explains why EEG has excellent temporal resolution—because the conduction itself is effectively instantaneous on the timescale of neuronal activity—but relatively poor spatial resolution. Only large, synchronous assemblies of neurons, particularly those in superficial cortical layers, contribute measurably to the scalp potential. Deeper or weakly synchronized sources contribute little to the recorded signal [1].

At a simplified level, the scalp potential $V(\mathrm{r}, t)$ at electrode position r can be approximated as the sum of contributions from discrete equivalent current dipoles:

$$V(\mathrm{r}, t) = \frac{1}{4\pi\sigma} \sum_{i=1}^{N} \frac{\mathrm{p}_i(t) \cdot (\mathrm{r} - \mathrm{r}_i)}{|\mathrm{r} - \mathrm{r}_i|^3} \tag{2.1}$$

where σ is the conductivity of the medium, $\mathrm{p}_i(t)$ is the dipole moment of source i, r_i is the dipole's location, and $|\mathrm{r} - \mathrm{r}_i|$ is the distance between the dipole and the electrode. This dipole summation model is particularly useful for conceptualizing how aligned cortical pyramidal cells act as macroscopic dipole generators.

A more general formulation of the *EEG forward problem* treats neuronal activity as a continuous distribution of current sources. The scalp potential $\phi(\mathrm{r}, t)$ can then be expressed as:

$$\phi(\mathrm{r}, t) = \int G(\mathrm{r}, \mathrm{r}')\, J(\mathrm{r}', t)\, d^3\mathrm{r}' \tag{2.2}$$

where $J(\mathrm{r}', t)$ is the primary current density at cortical location r' and time t, and $G(\mathrm{r}, \mathrm{r}')$ is the Green's function describing how currents at r' are transformed by the conductive properties of the head into potentials at electrode site r.

Together, these formulations highlight the dual nature of EEG modeling: while the dipole approximation provides an intuitive link between cortical geometry and EEG generation, the integral formulation emphasizes that the recorded potential reflects a spatially blurred, conductivity-weighted projection of distributed cortical sources [14,15].

2.4 Oscillatory Dynamics in the Brain

2.4.1 Foundations of Neural Oscillations

Neural oscillations reflect the synchronized activity of neuronal populations across multiple spatial and temporal scales. They arise from the recurrent interplay of excitatory pyramidal neurons and inhibitory interneurons, thalamo-cortical loops, and cortical network feedback. Unlike action potentials, which are unitary events, oscillations emerge primarily from the rhythmic modulation of postsynaptic potentials across cortical layers [14, 18, 20, 21]. Superficial layers (I–III) mediate long-range corticocortical communication and generate higher-frequency activity, while deep layers (V–VI) integrate thalamo-cortical input, producing lower-frequency oscillations. Each frequency band has distinct laminar origins and functional roles.

2.4.2 Frequency Bands and Functional Relevance

Delta (0.5–4 Hz): Delta oscillations dominate during deep non-REM sleep and states of cortical deafferentation. They are generated by layer V pyramidal neurons undergoing bistable up/down membrane potential transitions, with return currents in superficial layers, and reinforced by thalamo-cortical loops. Functional roles include cortical synchronization during sleep, gating of plasticity, and memory consolidation [22].

Theta (4–8 Hz): Theta rhythms are prominent in hippocampal and limbic circuits and appear in neocortex during working memory and attentional tasks. They arise from interactions between hippocampal pyramidal neurons and GABAergic interneurons, supported by septal pacemaker input. In cortex, theta reflects long-range corticocortical communication. Laminar origins include hippocampal CA1 dendrites and neocortical layer II/III. Functional roles include memory encoding, spatial navigation, and top-down control [18].

Alpha (8–13 Hz): Alpha oscillations are most pronounced in occipital cortex during eyes-closed rest but also appear in sensorimotor and parietal regions. They are generated by thalamo-cortical loops in which thalamic relay neurons rhythmically drive apical dendrites of pyramidal neurons in layers II/III via layer IV inputs. Alpha activity is thought to reflect functional inhibition of task-irrelevant regions and attentional gating. Laminar origin: layer IV input to pyramidal neurons in layers II/III [23].

Beta (13–30 Hz): Beta oscillations are prominent in sensorimotor cortex and basal ganglia–thalamocortical loops. In cortex, they arise from recurrent excitation and inhibition among layer V pyramidal neurons and fast-spiking interneurons. Beta activity is enhanced during motor maintenance and suppressed during movement initiation. Laminar origin: predominantly layer V pyramidal cells interacting with local interneurons. Functional roles include motor control, postural maintenance, and predictive coding [24].

Gamma (30–100 Hz): Gamma rhythms are generated via the pyramidal–interneuron gamma (PING) mechanism, where excitatory pyramidal neurons activate inhibitory interneurons, which then suppress pyramidal activity, creating oscillatory cycles. Gamma is strongest in superficial layers (II/III) due to dense local connectivity. Functional roles include feature binding, perceptual grouping, and attentional selection [20].

2.5 Invasive Techniques for Measuring Brain Activity

Invasive electrophysiological techniques offer the highest resolution for studying neural dynamics, as electrodes are placed directly into or onto neural tissue. Unlike non-invasive modalities, they bypass the skull and scalp, which typically attenuate and smear signals. This direct access allows for superior spatial precision and signal fidelity, though it comes at the cost of surgical intervention. Consequently, invasive approaches are predominantly employed in animal studies, human epilepsy monitoring, or brain–computer interface (BCI) research.

2.5.1 Intracellular Recording

Intracellular recording is one of the most fundamental approaches in neurophysiology, involving the insertion of a microelectrode directly into a single neuron to measure its membrane potential. With sharp electrodes or patch-clamp techniques, it is possible to monitor both fast action potentials and slower subthreshold postsynaptic potentials, as well as to isolate ionic currents flowing through individual channels [25]. This method has been instrumental in characterizing the basic biophysical mechanisms of excitability, synaptic transmission, and receptor pharmacology. However, it is highly invasive and technically demanding, usually restricted to in vitro brain slices or animal preparations. Because the procedure compromises cellular integrity, it is unsuitable for large-scale or chronic recordings in humans.

2.5.2 Extracellular Recording and Microelectrode Arrays

Extracellular recording places electrodes in the vicinity of neurons without penetrating the cell membrane, allowing the detection of voltage fluctuations in the surrounding extracellular space. These signals comprise two principal components: fast action potentials (spikes) generated by nearby neurons and slower local field potentials (LFPs), which reflect the summed synaptic and dendritic activity of larger neuronal populations [25]. Advances in microelectrode array (MEA) technology have enabled simultaneous recordings from dozens to thousands of neurons, supporting detailed analyses of ensemble coding, population dynamics, and large-scale network interactions.

Among extracellular recording technologies, penetrating microelectrode implants such as the Utah array represent a highly specialized and widely used class. A Utah array consists of a dense grid of silicon-based microelectrodes that penetrate the cortical surface, enabling the simultaneous sampling of activity from dozens to hundreds of neurons across multiple cortical columns. These arrays have been implanted chronically in both animal models and humans and have played a pivotal role in brain–computer interface (BCI) research. Notably, Utah arrays have enabled paralyzed individuals to control robotic limbs or computer cursors through decoded neuronal ensemble activity [29].

The primary advantages of extracellular recordings, including Utah arrays, are their relatively low invasiveness compared to intracellular methods, scalability, and suitability for long-term implantation. These properties make them particularly valuable for studying neural activity during naturalistic behaviors and for clinical and translational applications such as BCIs [26]. However, penetrating arrays also face important limitations. Chronic implantation can elicit tissue responses such as gliosis, which may degrade signal quality over time, and their spatial coverage is inherently limited relative to surface-based techniques such as electrocorticography (ECoG). Despite these challenges, penetrating extracellular microelectrode arrays remain

central to high-resolution, large-scale neural recording and represent a cornerstone of modern systems neuroscience and neuroengineering.

2.5.3 Depth Electrodes (Stereotactic EEG, sEEG)

Depth electrodes, or stereotactically implanted EEG (sEEG) electrodes, are used to access deep brain regions such as the hippocampus, amygdala, and insula. They are most commonly applied in clinical epilepsy monitoring, where they help localize seizure foci that cannot be reliably identified with surface EEG [27]. By recording both spikes and LFPs from deep structures, depth electrodes provide unique insights into large-scale oscillatory networks and their pathological alterations. Their major advantage lies in their ability to sample neural activity from regions inaccessible to non-invasive techniques, offering a bridge between local and global brain dynamics. However, the approach involves neurosurgical risks, including infection and hemorrhage, and is limited in spatial coverage since only the tissue surrounding the implanted trajectories is sampled.

2.5.4 Electrocorticography (ECoG)

Electrocorticography (ECoG) involves placing grids or strips of electrodes directly on the cortical surface beneath the dura mater. Because the skull is bypassed, ECoG signals are stronger and less distorted than scalp EEG, offering improved spatial resolution in the millimeter range while preserving millisecond temporal precision. Importantly, ECoG captures not only traditional oscillatory rhythms but also high-frequency broadband activity (70–200 Hz), which closely correlates with local neuronal firing [28]. These features make ECoG particularly powerful in both clinical and research contexts. Clinically, it is used for pre-surgical mapping of epileptogenic zones and eloquent cortical areas related to motor and language functions. In research, it has been leveraged for real-time BCIs, demonstrating higher reliability and bandwidth compared to scalp EEG. Its limitations, however, stem from the need for craniotomy and the fact that only the exposed cortical surface can be sampled, restricting both safety and coverage.

2.6 Non-Invasive Techniques for Measuring Brain Activity

Non-invasive techniques allow the study of brain activity without breaching the skull, making them safe, repeatable, and suitable for both clinical and research settings. Although they generally have lower spatial or temporal resolution compared to invasive methods, they provide access to large-scale neural dynamics in healthy populations. Each modality captures different aspects of neural activity, reflecting the trade-offs between safety, resolution, and accessibility [26]. Electroencephalography (EEG) is one of the most widely used non-invasive techniques for recording brain activity, offering excellent temporal resolution. In addition to EEG, several other non-invasive methods provide complementary information about neural dynamics, including magnetic, hemodynamic, and optical measures.

2.6.1 Magnetoencephalography (MEG)

Magnetoencephalography (MEG) records the tiny magnetic fields generated by postsynaptic currents in pyramidal neurons, using highly sensitive superconducting quantum interference devices (SQUIDs) or optically pumped magnetometers. Like EEG, MEG offers millisecond temporal resolution, but because magnetic fields are less distorted by the skull and scalp, it provides better spatial localization of sources, particularly for tangential cortical currents [28,30]. MEG is especially valuable in mapping functional connectivity and network

oscillations with high precision. Its clinical applications include pre-surgical localization of epileptic activity and functional mapping of sensory, motor, and language regions. Nevertheless, MEG systems are extremely costly, require magnetically shielded rooms, and remain largely confined to specialized research and clinical centers.

2.6.2 Functional Near-Infrared Spectroscopy(fNIRS)

Functional near-infrared spectroscopy (fNIRS) is an optical technique that measures changes in the concentration of oxygenated and deoxygenated hemoglobin in cortical tissue using near-infrared light. Because neuronal activity leads to localized increases in blood flow and oxygen consumption, fNIRS provides an indirect measure of neural activity via hemodynamic responses [31,32]. Its advantages include portability, relatively low cost, and tolerance for naturalistic settings, which makes it particularly useful for developmental and cognitive studies outside laboratory environments. However, its spatial resolution is limited to cortical surface regions, and its temporal resolution is constrained by the slow nature of hemodynamic responses, typically on the order of seconds. Compared to EEG and MEG, fNIRS cannot capture rapid neural dynamics, but its safety and practicality make it a promising tool for applied neuroscience, rehabilitation, and portable BCIs.

2.6.3 Functional Magnetic Resonance Imaging (fMRI)

Functional magnetic resonance imaging (fMRI) measures brain activity by detecting blood-oxygen-level-dependent (BOLD) signals, which reflect changes in the balance of oxygenated and deoxygenated hemoglobin associated with neural activation [33,34]. fMRI provides excellent spatial resolution, often at the millimeter scale, and has been pivotal in mapping large-scale functional networks and investigating the organization of sensory, motor, and cognitive systems. However, its temporal resolution is limited to seconds, given its reliance on hemodynamic responses, making it unsuitable for tracking rapid neuronal oscillations. Despite this limitation, fMRI remains the dominant tool in systems neuroscience due to its ability to provide whole-brain coverage non-invasively. Its disadvantages include high cost, lack of portability, and sensitivity to motion artifacts, which constrain its use in certain populations and experimental designs.

In summary, non-invasive techniques including EEG, MEG, fNIRS, and fMRI provide complementary insights into brain activity by balancing safety, accessibility, and resolution. EEG and MEG capture fast neural dynamics with high temporal precision, whereas fNIRS and fMRI measure slower hemodynamic responses with superior spatial coverage. Although these methods cannot resolve single-neuron activity, they are indispensable for investigating large-scale brain networks, enabling longitudinal studies in healthy populations and supporting clinical and brain–computer interface applications.

Chapter 3

Brain Computer Interfacing Using Motor Imagery EEG

3.1 Introduction

Brain-Computer Interfaces (BCIs) are systems that enable direct communication between the brain and external devices, bypassing conventional neuromuscular pathways. The primary goal of a BCI is to translate neural activity into actionable commands, which can be used to control assistive technologies, robotic systems, or computer applications. BCIs have applications in clinical settings, such as restoring motor function in paralyzed patients or providing communication channels for individuals with severe neuromuscular disorders, as well as non-clinical domains, including gaming, neurofeedback, and human-computer interaction [35,36]

Among the various neural recording modalities, EEG is the most commonly used for BCI systems due to its high temporal resolution, non-invasiveness, portability, and relative affordability. Although EEG has lower spatial resolution compared to techniques such as MEG or ECoG, it provides sufficient information to decode motor intentions, attention, or cognitive states in real-time, making it ideal for practical BCI implementations [37].

3.2 Brain Computer Interfacing Paradigms

3.2.1 P300 Event-Related Potential (ERP)

The P300 paradigm exploits an oddball task, in which infrequent target stimuli are embedded within frequent non-target stimuli. Approximately 300 ms after presentation of a target stimulus, a positive deflection in EEG voltage appears over parietal sites (Pz). This P300 response can be detected and used for applications such as spelling systems, selection interfaces, and communication aids. It offers high accuracy even with minimal training, requires the user's attention and stimulus discrimination, and is typically recorded using parietal electrodes [38,39].

3.2.2 Steady-State Visual Evoked Potentials (SSVEP)

SSVEP BCIs exploit the brain's natural response to flickering visual stimuli. When a user focuses on a light source flickering at a specific frequency (e.g., 10 Hz), the EEG over the occipital cortex exhibits corresponding frequency peaks. These BCIs can achieve high information transfer rates due to the strong frequency-tagged response and minimal user training. Signals are recorded primarily from occipital electrodes (O1, O2, Oz), and frequency-domain features such as power spectral density or canonical correlation analysis are used for classification, enabling robust multi-choice selection interfaces [40,41].

3.2.3 Slow Cortical Potentials (SCPs)

Slow Cortical Potentials (SCPs) are slow shifts in the EEG baseline, typically lasting from several hundred milliseconds to several seconds, reflecting changes in cortical excitability. Negative SCPs indicate increased cortical activation, while positive SCPs correspond to decreased excitability. SCP-based BCIs allow users to voluntarily modulate these potentials through neurofeedback training, enabling communication, control of external devices, and rehabilitation applications. Signals are typically recorded from central or frontocentral electrodes, and classification relies on detecting the direction and amplitude of the potential shifts [42,43].

3.2.4 Error-Related Potentials (ErrPs)

Error-Related Potentials (ErrPs) are EEG components elicited when a user detects an error, either in their own actions or in a system's response. Typically appearing as a negative–positive deflection over frontocentral electrodes within 100–300 ms after the error, ErrPs can be harnessed in BCIs to automatically correct mistakes or adapt system behavior. Their detection enables rapid error compensation in applications such as cursor control, robotic assistance, and adaptive interfaces, enhancing reliability without requiring explicit user intervention [44,45].

3.2.5 Motor Imagery (MI)

Motor imagery involves the mental rehearsal of movement without actual execution, such as imagining movements of the left hand, right hand, or feet, which produces characteristic modulations in sensorimotor rhythms (SMR), primarily in the Mu(8–13 Hz) and Beta (13–30 Hz) bands over the motor cortex. These changes manifest as event-related desynchronization (ERD) during imagined movement and event-related synchronization (ERS) after imagination ends. MI paradigms are extensively used in rehabilitation BCIs to restore motor function in paralyzed patients or to control prosthetic devices, with EEG channels over the sensorimotor cortex (C3, C4, Cz) commonly employed and temporal and spectral features extracted to detect ERD/ERS patterns, while performance can be enhanced through user training and feedback [37,46].

3.2.5.1 Network Level Functional Connectivity

Motor imagery (MI) refers to the mental rehearsal of a movement without any overt muscular execution. Neuroimaging and electrophysiological studies have demonstrated that imagining a movement activates neural substrates that substantially overlap with those engaged during actual motor performance, including the primary motor cortex (M1), premotor cortex, supplementary motor area (SMA), and parietal regions [47,48]. This overlap is often attributed to the mirror neuron system, which underpins both action observation and internal simulation of motor acts.

Motor imagery (MI) is not merely the activation of isolated motor areas but a complex, distributed process that relies on dynamic, network-level interactions to simulate movement without overt execution. Central to this process is the functional interplay between cortical and subcortical regions, particularly the supplementary motor area (SMA) and primary motor cortex (M1). EEG and fMRI studies consistently show that the SMA plays a regulatory role over M1, potentially involving inhibitory mechanisms that help suppress overt motor output during mental rehearsal [49,50], though the precise neurophysiological pathways — whether

direct cortico-cortical projections or mediated via basal ganglia-thalamo-cortical loops — remain under active investigation.

Beyond cortico-cortical interactions, thalamo-cortical circuits critically modulate sensorimotor dynamics during MI. Specifically, reciprocal connections between the ventral lateral (VL) thalamic nucleus and cortical motor areas (including M1 and SMA) are implicated in gating sensorimotor signals and dynamically regulating cortical oscillations — particularly in the Mu(8–13 Hz) and Beta (13–30 Hz) frequency bands — which reflect motor preparation and suppression [51,52]. The thalamic reticular nucleus may further contribute to inhibitory control through GABAergic projections, reinforcing the brain's ability to maintain motor inhibition during mental simulation.

Crucially, functional connectivity during MI is often frequency-specific. Metrics such as phase-locking value, coherence, or imaginary coherence in the Mu and Beta bands reveal how sensorimotor, parietal, and frontal regions synchronize or desynchronize to support the integration of motor planning, body schema representation, and attentional control [53]. These oscillatory couplings evolve dynamically over the course of imagined movement, underscoring the need for neuroimaging tools with high temporal resolution.

In addition to SMA-M1-thalamic circuits, MI consistently recruits a broader network: premotor areas encode action sequences and intentions; parietal regions — especially the superior and inferior parietal lobules — represent spatial and somatosensory aspects of movement and body schema [54,55]; the cerebellum supports internal modeling, timing, and coordination [56]; and prefrontal regions contribute to working memory and attentional control, particularly during complex or multi-step imagery [57]. Together, these findings confirm that MI is a multi-regional, systems-level process.

This distributed nature has direct methodological implications. Individual variability in cortical anatomy, functional organization, and preferred imagery modality (kinesthetic vs. visual) leads to substantial differences in the spatial distribution and amplitude of neural activity across participants [59]. Such variability not only affects the localization of event-related desynchronization/synchronization (ERD/ERS) but may also alter the topology and strength of functional networks — a critical consideration for brain-computer interface (BCI) calibration and group-level neuroimaging analyses.

High-density EEG systems (e.g., 64-channel montages) offer a practical and effective compromise: they provide sufficient spatial sampling to capture activity across frontal, parietal, and sensorimotor regions while preserving the millisecond temporal resolution necessary to track oscillatory coupling and event-related dynamics. Although they lack the spatial precision of fMRI, their ability to resolve fast-changing network interactions makes them uniquely suited for studying the real-time functional connectivity underlying MI.

3.2.5.2 Event-Related Desynchronization and Synchronization (ERD/ERS)

A defining feature of motor imagery in EEG is its modulation of oscillatory activity in the sensorimotor cortex, expressed through event-related desynchronization (ERD) and event-related synchronization (ERS). ERD refers to the suppression of rhythmic activity, particularly in the Mu(8–13 Hz) and Beta (13–30 Hz) frequency bands, during imagined movement. This desynchronization reflects the recruitment and increased excitability of neuronal populations, reducing the synchrony of ongoing oscillations. In contrast, ERS represents a rebound in

rhythmic power following the termination of imagery, often observed as a Beta rebound, and is thought to signify cortical deactivation, recovery, and inhibition of motor circuits [52,59].

These ERD/ERS patterns serve as the fundamental neurophysiological signals that enable MI-based BCIs, allowing users to modulate their brain rhythms voluntarily to communicate intentions or control external devices.

3.2.5.2.1 Kinesthetic vs. Visual–Motor Imagery

An important distinction in motor imagery research lies between kinesthetic motor imagery (KMI) and visual motor imagery (VMI), which differ in their underlying neurophysiological dynamics. Both modalities engage the Mu (8–13 Hz) and Beta (13–30 Hz) rhythms, yet their modulation patterns are distinct.

According to Neuper et al. [60], KMI typically elicits stronger event-related desynchronization (ERD) in the mu and Beta frequency bands over contralateral sensorimotor areas, reflecting enhanced recruitment of motor-related cortical circuits involved in movement preparation and execution. This desynchronization is frequently followed by a Beta rebound, or event-related synchronization (ERS), once the imagery ceases, indicating cortical deactivation and recovery.

By contrast, VMI generally produces weaker sensorimotor ERD and is instead characterized by synchronization in occipital and parietal regions, consistent with the reliance on visual imagery and spatial representation networks. These findings highlight that VMI recruits cortical processes associated with visual perception and mental simulation of observed movements, rather than the kinesthetic aspects of motor execution.

In their experimental paradigm, Neuper and colleagues instructed participants to perform either KMI or VMI of right-hand movements. In the KMI condition, participants imagined clenching a ball with their right hand, focusing on the internal, kinesthetic sensation of movement while avoiding overt muscle activation. In the VMI condition, subjects were asked to visualize right-hand movements as though watching a mental "video" of the previously observed "alien" hand displayed on a monitor. The results clearly demonstrated that KMI engages contralateral sensorimotor regions more strongly, whereas VMI predominantly activates occipital and parietal regions, thereby underscoring the task-specific recruitment of cortical networks.

To further illustrate the neurophysiological distinctions between kinesthetic and visual motor imagery, Figure. 6 presents time–frequency maps of event-related desynchronization (ERD) and synchronization (ERS) across key cortical regions during KMI and VMI tasks from Neuper et al. [60]. These maps, derived from EEG recordings in three representative subjects (p16, X4, X8), reveal the spatially and temporally specific modulation of Mu and Beta rhythms over sensorimotor (C3, C4) and parieto-occipital sites. KMI elicits pronounced ERD (red, $P < 0.01$) over contralateral sensorimotor cortex, followed by post-imagery Beta rebound (blue, $P > 0.01$), while VMI shows attenuated sensorimotor modulation and instead features stronger synchronization in posterior regions — aligning with its reliance on visual-spatial processing networks. This visualization not only confirms the dissociable cortical signatures of the two imagery modalities but also underscores the importance of modality-specific protocol design in brain–computer interface (BCI) and neurorehabilitation applications.

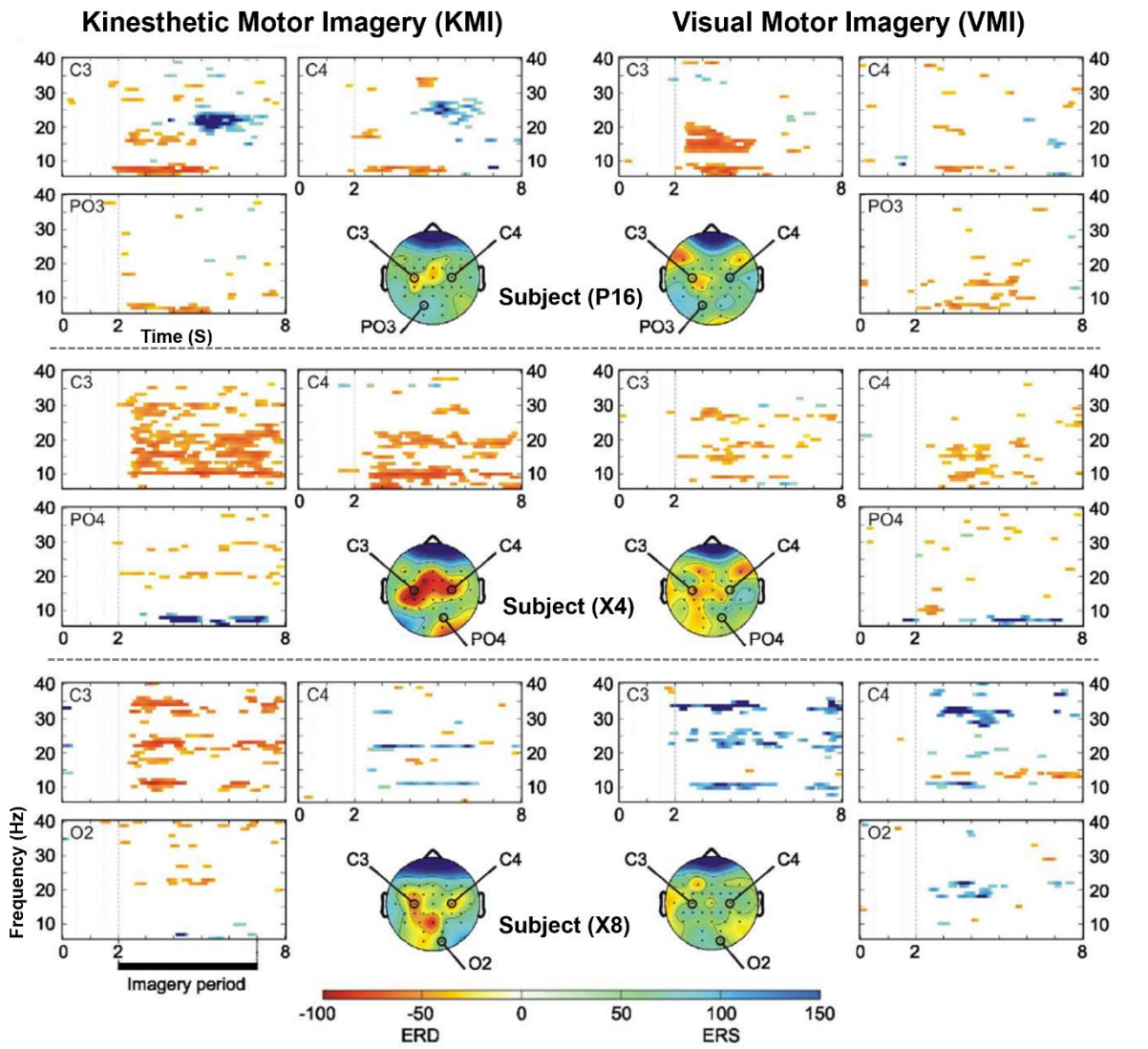

Figure 6: ERD/ERS time–frequency maps during kinesthetic motor imagery (KMI) and visual–motor imagery (VMI) for three subjects (p16, X4, X8). Maps are shown for left (C3) and right (C4) sensorimotor sites and one parieto-occipital site. Red indicates significant power decreases (ERD, $P < 0.01$), and blue indicates significant power increases (ERS, $P > 0.01$) [60].

3.2.5.2.1 Beta Rebound

Oscillatory activity in the Beta frequency band (~13–30 Hz) is a hallmark of the human sensorimotor system. Since the pioneering work of Jasper and Penfield in 1949 [61], it has been established that motor-related cortical areas generate Beta oscillations, typically peaking around 20 Hz, which are suppressed — or desynchronized — prior to and during voluntary movement. This event-related desynchronization (ERD) localizes to cortical regions representing the activated body part and has been robustly confirmed using electrocorticography [62,63] and magnetoencephalography [64,65]. Beta rhythms predominantly originate from the anterior bank of the central sulcus, whereas the Mu rhythm (~10 Hz) shows stronger contributions from the postcentral gyrus [64]. Critically, Beta desynchronization is not limited to overt movement: it also occurs during motor planning and motor imagery [66], underscoring shared network engagement.

Following movement termination, Beta activity rapidly rebounds — typically within ~1 second — often exceeding baseline levels in brief, localized bursts of synchronization. This phenomenon, termed the post-movement Beta rebound [67], emerges over the same cortical territories that exhibited ERD. PMBR is not exclusive to motor acts; it also follows somatosensory stimulation, such as median nerve stimulation [64,68]. Notably, this somatosensory-induced rebound is suppressed when subjects concurrently perform motor tasks or MI [68, 69], reinforcing the notion that MI and execution recruit overlapping sensorimotor networks responsible for Beta rhythm modulation [69].

The functional significance of PMBR remains actively debated. While one prominent interpretation posits that it reflects active inhibition or cortical deactivation following movement [65, 67], alternative hypotheses suggest it may encode somatosensory reafference [70], prediction error signaling [71], or network-level resetting to stabilize cortical dynamics [72]. Supporting the inhibition hypothesis, transcranial magnetic stimulation (TMS) studies show reduced motor cortex excitability during the PMBR window — observed after both finger movements [73] and median nerve stimulation [74]. This suppression of corticospinal excitability aligns with the view that Beta synchronization serves to "hold" the motor system in a stable, inhibited state [75].

The topography and amplitude of PMBR are strongly effector-dependent. Pfurtscheller et al. [67] demonstrated distinct spatial patterns of post-movement Beta synchronization following hand, foot, and tongue imagery. Foot imagery, in particular, elicits a pronounced Beta rebound maximal at the vertex (electrode Cz). This may be attributed to the anatomical proximity of the foot motor representation to the supplementary motor area (SMA), facilitating synchronized Beta oscillations across these regions [76,77]. The resulting signal amplification at midline electrodes likely reflects synchronized inhibition across an extended cortical territory encompassing both SMA and primary motor cortex leg areas [52, 78], suggesting PMBR may serve as a mechanism for resetting overlapping motor networks.

Functionally, PMBR is increasingly interpreted as a correlate of post-movement inhibitory resetting — a return of the motor system to baseline after activation. Its amplitude, latency, and spatial distribution are modulated by task type (imagery vs. execution), effector, and cognitive state. These properties make PMBR a valuable biomarker in both basic and applied neuroscience. In brain–computer interfaces (BCIs), PMBR can serve as a robust "movement termination" signal, improving decoding accuracy and user control [79]. In neurorehabilitation, attenuated or delayed PMBR has been associated with impaired motor recovery after stroke [80], suggesting its potential as a prognostic marker or target for neuromodulation therapies.

3.2.5 Movement-Related Potentials and Self-Paced Motor Imagery

Movement-related potentials (MRPs), also known as Bereitschaftspotential, are slow EEG potentials that reflect the cortical dynamics of voluntary movement preparation and initiation. They comprise three main components: the readiness potential (RP), the negative slope (NS), and the motor potential (MP), with the contingent negative variation (CNV) also contributing in cue-based paradigms.

The readiness potential (RP) is the earliest component, beginning approximately 1–2 seconds before movement onset. It reflects the gradual buildup of neural activity associated with motor intention and planning. RP is classically divided into two phases. The early RP, which emerges 1.5–2 seconds prior to movement, is maximal at the midline electrode Cz and originates

primarily in the supplementary motor area (SMA) and pre-SMA [81,82]. The late RP, beginning around 500–200 ms before movement, represents more focused motor preparation. Its terminal portion, the negative slope (NS), steepens dramatically in the final ~200 ms and becomes strongly lateralized over contralateral motor regions (C3/C4), reflecting recruitment of the primary motor cortex (M1) for movement execution [76].

The motor potential (MP) is a sharp negativity that peaks at the moment of movement onset. Maximal over contralateral central electrodes, the MP is generated primarily by M1 and reflects corticospinal output responsible for triggering muscle contraction [82].

During motor execution, all three MRP components (early RP, late RP/NS, and MP) are expressed. In motor imagery, this structure is selectively altered. The early RP is reliably present, reflecting SMA and premotor engagement during imagined movement preparation [83]. In contrast, the late RP/NS is attenuated or absent, indicating reduced M1 involvement and likely active suppression of corticospinal output [84,85]. Similarly, the MP does not typically emerge during motor imagery, confirming that motor simulation terminates before efferent activation.

Interestingly, in some individuals — particularly those engaging in vivid kinesthetic motor imagery — weak NS-like lateralized activity can be detected, suggesting partial but subthreshold recruitment of M1 without overt execution [59]. This interindividual variability underscores the influence of imagery modality and neural efficiency in shaping MRP expression.

Overall, MRPs demonstrate that motor imagery activates early preparatory circuits but inhibits late execution-related activity. This suppression likely involves top-down regulation from SMA and thalamo-cortical loops to prevent overt motor output. As such, MRPs provide a robust electrophysiological signature distinguishing imagery from execution. In asynchronous brain–computer interfaces (BCIs), the early RP serves as a natural, self-paced intention signal, while the absence of MP confirms that no actual movement occurred [59]. In neurorehabilitation, repeated motor imagery combined with MRP-based neurofeedback may reinforce preparatory motor network activation and promote cortical reorganization in patients with motor impairments [86].

3.2.5.4 Contingent Negative Variation and Cue-Based Motor

The Contingent Negative Variation (CNV) is a slow cortical potential that emerges during cue–target paradigms, appearing in the interval between a warning stimulus (S1) and an imperative stimulus (S2) [88]. The CNV is classically divided into two phases: an early CNV, reflecting orienting and attentional processes immediately after the warning cue, and a late CNV, which builds up prior to the imperative cue and indexes anticipatory motor preparation [89].

In the context of motor imagery, the CNV is highly relevant because most experimental paradigms and open-access datasets, such as the PhysioNet EEG Motor Movement/Imagery Database [90,91], employ cue-based designs. When participants are instructed to imagine a movement following an auditory or visual cue, the CNV reliably appears, although with reduced amplitude compared to actual execution [92]. This attenuation suggests that while attentional orienting and preparatory activity are engaged during motor imagery, corticospinal output is suppressed.

The late CNV is particularly important for motor imagery tasks, as it overlaps functionally with the early readiness potential (RP) recorded in self-paced paradigms. Both reflect activity within supplementary motor and premotor areas, associated with motor planning without overt execution [82, 87]. Because the CNV is locked to cue presentation, it provides a temporally precise marker of motor preparation during imagery, making it useful for event-related potential (ERP)-based brain–computer interfaces [52,93].

Thus, the CNV represents a complementary neural signature to MRPs in motor imagery research. While MRPs dominate in self-paced designs, CNV is the hallmark of cue-based tasks. Its consistent expression in datasets such as PhysioNet underlines its value for studying the temporal dynamics of motor imagery and developing ERP-based control signals for rehabilitation and BCI applications.

3.2.5.5 Cognitive Dimensions of Motor Imagery

Mental imagery is a multidimensional construct. While vividness is often emphasized as the most salient aspect, a complete account of imagery in the motor domain also requires considering other characteristics such as duration, controllability, perspective, modality, accuracy, emotional valence, and neural engagement. Each of these dimensions contributes to the effectiveness of motor imagery (MI) in skill acquisition, performance enhancement, rehabilitation, and BCI systems.

3.2.5.5.1 Vividness

Imagery vividness refers to the clarity and sensory richness with which an individual can imagine a movement. It is typically assessed with self-report questionnaires such as the Movement Imagery Questionnaire [94] or the Vividness of Visual Imagery Questionnaire [95]. High vividness is associated with stronger cortical activation in motor-related regions, including the supplementary motor area (SMA), premotor cortex, and primary motor cortex (M1) [54, 55]. Electrophysiologically, vivid imagery is linked to pronounced event-related desynchronization (ERD) in the Mu (8–12 Hz) and Beta (13–30 Hz) bands, indicating robust recruitment of sensorimotor networks [52, 59].

At the extremes of the vividness spectrum lie hyperphantasia [96] and aphantasia [97]. Recent findings show that aphantasia extends beyond visual imagery to include deficits in motor imagery, with individuals showing reduced motor cortex excitability during MI, as reflected in absent or diminished motor-evoked potentials [98, 99].

3.2.5.5.2 Duration

Imagery duration refers to how long an individual can sustain a mental image. Mental images typically decay within seconds unless actively maintained [100]. Individuals with high imagery vividness can often sustain imagery longer, which correlates with more stable activation in sensory and motor cortices [101, 102]. In the motor domain, sustaining imagery over time has been shown to enhance corticospinal excitability and produce stronger ERD in EEG, reflecting deeper engagement of sensorimotor networks [59]. Thus, duration is an important factor for the effectiveness of imagery practice in sports and rehabilitation.

3.2.5.5.3 Controllability

Controllability refers to the ability to manipulate and transform mental images, such as rotating a movement, changing speed, or shifting sequence order. In motor imagery, high controllability allows for flexible rehearsal of complex skills and adaptations to novel scenarios [103]. Athletes with high controllability can, for example, mentally simulate a gymnastics routine forward and backward or practice movement corrections mentally before executing them physically.

3.2.5.5.4 Perspective

Imagery can be experienced from an internal perspective (first-person view, as if performing the action oneself) or an external perspective (third-person view, observing oneself as if from outside). Internal perspective imagery is more strongly associated with kinaesthetic sensations and greater activation of motor cortex regions, making it particularly effective for MI [104]. By contrast, external perspective imagery recruits more visual cortical areas and may be advantageous when focusing on form, posture, or aesthetic aspects of movement [105].

3.2.5.5.5 Modality (Sensory Components)

Motor imagery encompasses multiple sensory modalities that contribute to the richness and effectiveness of the imagined experience. These modalities include visual imagery, in which the individual generates a mental representation of the appearance and spatial characteristics of the movement; kinaesthetic imagery, which involves the internal sensation of performing the action, such as muscle tension, joint position, and bodily effort; auditory imagery, where sounds associated with the movement (e.g., the impact of a tennis ball on a racket) are mentally simulated; and tactile imagery, which relates to the perception of textures or contact forces during action. Among these modalities, kinaesthetic imagery has been consistently identified as the most critical for motor learning and rehabilitation, given its stronger engagement of sensorimotor networks compared to visual imagery alone [60]. This kinaesthetic component is thought to more closely mirror the neural processes underlying actual motor execution, thereby facilitating motor skill acquisition and recovery. In line with this distinction, the Movement Imagery Questionnaire–3 [94] explicitly separates visual and kinaesthetic imagery dimensions, allowing for a more precise assessment of an individual's imagery profile.

3.2.5.5.6 Accuracy and Fidelity

Accuracy refers to the degree to which imagined movements match real movements in timing and kinematics. A well-documented finding is temporal congruence: in skilled imagers, the duration of imagined movements closely matches physical execution [106]. Such fidelity is critical in sports training and rehabilitation, where motor imagery is intended to replicate real performance dynamics. Disruptions in temporal congruence (e.g., in neurological conditions) can reduce the efficacy of MI.

3.2.5.5.7 Emotional Valence and Intensity

Imagery is not purely cognitive—it can be emotionally charged. The emotional content of imagery influences physiological arousal and neural engagement [107]. In motor contexts, positive and motivating imagery can enhance performance, while negative imagery may interfere with execution or increase anxiety. Thus, tailoring emotional valence is a practical consideration in applied MI interventions.

3.2.5.5.8 Neural Engagement

Finally, imagery dimensions are grounded in variability in neural engagement. Neuroimaging and neurophysiological studies reveal that individual differences in imagery vividness, duration, and modality preference are mirrored in the degree of motor system recruitment [54, 55]. For instance, phantasics show clear corticospinal facilitation during MI, whereas individuals with aphantasia may not [98].

3.2.5.6 Motor Imagery Illiteracy

Despite training, a subset of individuals exhibits persistent difficulty in performing effective MI, a phenomenon referred to as "MI illiteracy." Estimates suggest that approximately 10–30% of healthy adults fall into this category [108, 109], which has significant implications for BCI and neurorehabilitation applications. MI illiteracy may arise from poor kinaesthetic awareness, reduced connectivity between SMA, premotor areas, and M1, or inadequate top-down control for suppressing irrelevant cortical activity during MI [59]. Strategies to mitigate MI illiteracy include guided imagery training, multimodal feedback (visual, auditory, or haptic), and task adaptation tailored to individual ability levels [110, 111]. Early identification of MI-impaired individuals is therefore critical to optimize interventions and improve engagement with MI-based therapies.

Individual differences in MI ability are substantial, and understanding these differences is essential for effective application of MI in rehabilitation, skill acquisition, and BCI systems.

3.2.6 EEG Acquisition and Signal Processing

EEG signals are recorded from the scalp using electrodes and require careful acquisition to ensure signal quality. Subsequent signal processing, including filtering, artifact removal, and feature extraction, is essential to isolate meaningful neural activity for analysis and BCI applications.

3.2.6.1 EEG Hardware and Setup

3.2.6.1.1 EEG Electrode Montage Systems

EEG recordings rely on standardized electrode placement systems to ensure consistent spatial coverage of the scalp. The most used systems are the 10–20, 10–10, and 10–5 montages, which provide increasing spatial resolution[112].

The 10–20 System: is the classical EEG montage, widely used in clinical and research settings [113]. It is based on proportional distances between anatomical landmarks on the scalp, including the nasion, inion, and preauricular points. This system consists of 21 electrodes, labeled according to their approximate brain region (e.g., Fp, F, C, P, O, T) and hemispheric location. Electrodes on the right hemisphere are numbered with even numbers, electrodes on the left hemisphere with odd numbers, and electrodes along the midline with "z" labels. This montage provides sufficient coverage for general EEG analysis but has limited spatial resolution for high-density studies [53].

The 10–10 System: extends the 10–20 montage by adding 53 additional electrodes, resulting in a total of 74 positions [114]. By reducing the inter-electrode distance to 10% of the nasion-inion or preauricular distances, this montage offers denser scalp coverage. Due to its optimal balance of coverage and practicality, the 10–10 system is now the de facto standard in most BCI and cognitive neuroscience research using 64- or 128-channel systems [115].
The 10–5 System: further increases electrode density and is commonly used in high-density

EEG recordings (e.g., 128 or 256 channels) [116]. In addition to the 10–10 electrodes, it includes extra positions indicated as dots or open circles in scalp maps. This montage allows for precise localization of cortical activity and better capture of spatially distributed EEG patterns. The labeling convention follows the same rules as the 10–20 system. Figure.7 illustrates the standard EEG montages and highlights the spatial adjacency and progressive increase in electrode density from the 10–20 to the 10–5 system. and it shows the 10–5 electrode positions on a realistic head model from different perspectives—lateral, anterior, and posterior—providing an intuitive understanding of how electrodes uniformly cover the scalp surface [117].

3.2.6.1.2 EEG Hardware and Setup Considerations

EEG hardware setup involves several critical considerations to ensure high-quality signal acquisition. Electrodes may be wet, using conductive gel or paste, or dry, which are faster to set up but generally noisier. Active electrodes incorporate preamplifiers at the electrode site to improve signal quality, whereas passive electrodes rely entirely on the main amplifier. Amplifiers used in EEG systems must have high input impedance and low noise to accurately capture microvolt-level brain signals. The number of channels also affects spatial resolution: low-density systems (e.g., 21 electrodes) are sufficient for general monitoring, while high-density systems (64–256 electrodes) enable advanced analyses, albeit with longer setup times, increased participant discomfort, and larger data volumes. Finally, practical setup considerations—including maintaining electrode impedance below 10 kΩ, proper skin preparation, and correct placement of reference and ground electrodes—are essential to reduce noise and ensure stable recordings [118].

3.2.6.2. EEG Preprocessing

3.2.6.2.1 EEG Filtering

EEG signals contain a mixture of neural activity and various sources of noise, including line interference, muscle activity, and slow drifts. Filtering is an essential preprocessing step to isolate the frequency bands of interest, remove unwanted components, and improve the signal-to-noise ratio. Commonly used filters in EEG processing include band-pass, low-pass, high-pass, and notch filters. Band-pass filters are typically employed to extract task-relevant rhythms such as Mu (8–13 Hz) and Beta (13–30 Hz) oscillations associated with motor imagery, while notch filters (e.g., at 50 or 60 Hz) remove line noise.

Among the different types of filters, the Butterworth filter is widely used due to its maximally flat frequency response in the passband and smooth attenuation in the stopband. Its transfer function $H(s)$ in the analog domain is given by:

$$|H(j\omega)|^2 = \frac{1}{1+(\omega/\omega_c)^{2n}} \tag{3.1}$$

where ω_c is the cutoff frequency, nnn is the filter order, and ω is the angular frequency. The digital implementation of the Butterworth filter can be applied using standard signal processing libraries, and the *filtfilt* function is often used for offline EEG analysis to perform zero-phase filtering, applying the filter forward and backward to eliminate phase distortions. This preserves the temporal features of the EEG signal, which is particularly important for event-related potentials and motor imagery studies. By carefully selecting the cutoff frequencies and

filter order, Butterworth band-pass filters allow researchers to retain the relevant neural oscillations while attenuating noise, forming a critical step in EEG preprocessing pipelines [119, 120].

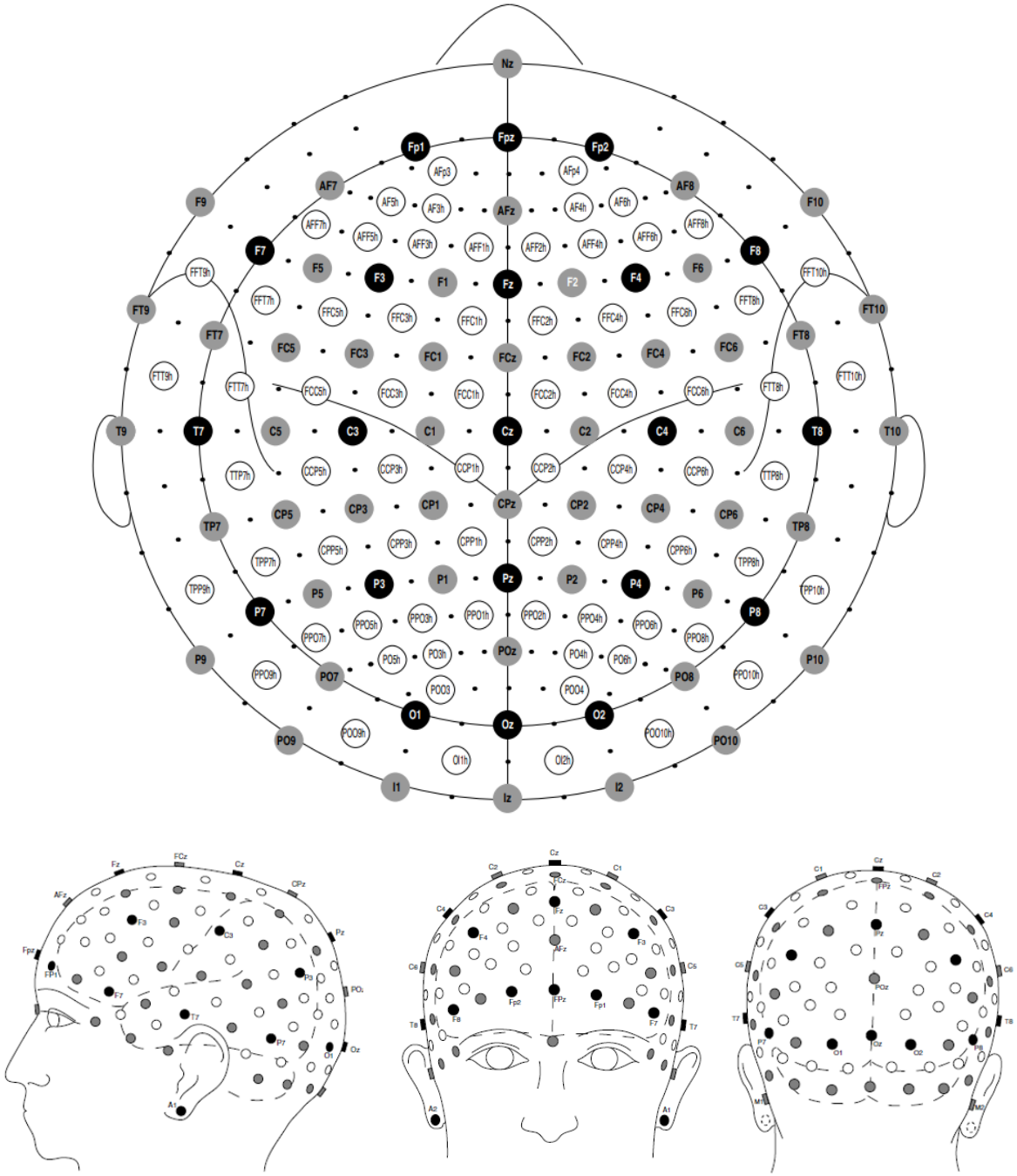

Figure 7: Comparison of EEG electrode density and placement across standard montages. Upper panel: Electrode density across three common EEG montages. The international 10–20 system consists of 21 electrodes (black circles). The extended 10–10 system adds 53 intermediate electrodes for a total of 74 (grey circles). The 10–5 montage further increases spatial resolution to a 128-channel system, with additional positions indicated by black dots and selected electrodes highlighted as open circles. Lower panel: Realistic head model illustrating electrode positions in the 10–5 montage. Left: lateral view of the left hemisphere; middle: anterior view; right: posterior view [117].

3.2.6.2.2 Artifact Rejection

EEG signals are inherently prone to contamination from various artifacts, including eye movements and blinks (EOG), muscle activity (EMG), cardiac signals (ECG), and environmental noise such as line interference. These artifacts can obscure neural activity, particularly the subtle oscillations associated with motor imagery, and must be addressed during preprocessing. Artifact rejection refers to the techniques used to remove or attenuate these unwanted signals. Approaches range from manual inspection, where contaminated segments are discarded, to automatic thresholding, which flags epochs exceeding predefined voltage or frequency limits. Advanced methods include Independent Component Analysis (ICA), regression- or wavelet-based techniques, and fully automated algorithms such as FASTER (Fully Automated Statistical Thresholding for EEG artifact Rejection). FASTER applies statistical thresholds across multiple dimensions—channels, epochs, and independent components—to detect and remove artifacts without human intervention, making it particularly useful for large EEG datasets or high-density recordings [121].

3.2.6.2.3 Baseline Correction

EEG signals are often subject to slow drifts and inter-trial variability that can obscure task-related activity. Baseline correction addresses this by adjusting each trial relative to a predefined reference interval, usually taken from a period immediately before stimulus onset or motor imagery cue. The average voltage in this baseline window is computed and subtracted from the subsequent signal:

$$X'(t) = X(t) - \overline{X}_{baseline} \tag{3.2}$$

where $X(t)$is the raw signal at time ttt and $\overline{X}_{baseline}$ is the mean over the baseline period. This procedure centers the data around zero, removing slow drifts and ensuring that changes in the EEG reflect task-related activity rather than global shifts in signal amplitude.

In addition to simple subtraction, relative baseline correction is also widely used:

$$X_{\text{rel}}(t) = \frac{X(t) - \overline{X}_{baseline}}{\overline{X}_{baseline}} \tag{3.3}$$

which expresses deviations as percentages relative to the baseline.

For time–frequency analysis, normalization is often applied to spectral power. If $P(f,t)$ is the power at frequency f and time t, and $\overline{P}_{baseline}(f)$is the mean baseline power at the same frequency, then:

$$P_{\text{rel}}(f,t) = \frac{P(f,t) - \overline{P}_{baseline}(f)}{\overline{P}_{baseline}(f)} \times 100\% \tag{3.4}$$

To stabilize variance and emphasize proportional changes, many studies apply a log-based baseline normalization:

$$P_{\text{norm}}(f,t) = \log\big(P(f,t)\big) - \log\left(\overline{P}_{baseline}(f)\right) = \log\left(\frac{P(f,t)}{\overline{P}_{baseline}(f)}\right) \tag{3.5}$$

This method highlights relative changes in oscillatory power while minimizing scale differences across frequencies and subjects, making it particularly effective for detecting event-related desynchronization/synchronization (ERD/ERS) patterns in motor imagery tasks [53, 122].

3.2.6.2.4 Referencing Schemes

Electroencephalography (EEG) signals are measured as the difference in voltage between an active electrode and a reference, making the choice of reference critical for accurate interpretation. Common referencing schemes include single-electrode references, average reference, linked-mastoids or ears, Laplacian (current source density, CSD), and the Reference Electrode Standardization Technique (REST). A single-electrode reference, such as a mastoid or earlobe, is simple but can bias the signal toward the reference site, while linked-mastoid referencing reduces lateralized bias by averaging two sites.

Average referencing, widely used in high-density EEG, re-references each electrode to the mean of all electrodes, calculated as

$$V_{new}(i) = V(i) - \frac{1}{N}\sum_{j=1}^{N} V(j) \tag{3.6}$$

where $V(i)$ is the original potential at electrode i and N is the total number of electrodes, helping to distribute the reference influence evenly [14, 123]. Laplacian or CSD referencing improves spatial resolution by referencing each electrode to the average of its neighbors, reducing the effects of volume conduction [124]. REST is a computational method that estimates a neutral reference at infinity, minimizing reference-dependent bias [125]. Since no reference is truly neutral, offline re-referencing is commonly applied to optimize signal quality according to the experimental goals.

3.2.6.2.5 Normalization

EEG signals exhibit high variability due to physiological differences, electrode placement, and recording conditions. To make signals or extracted features comparable across channels, trials, or subjects, normalization techniques are applied. These methods adjust the scale or distribution of data, improving statistical analyses and machine learning performance. The three most common methods are Z-score normalization, Min–Max normalization, and log-transform normalization.

Log-transform Normalization

EEG spectral power often spans multiple orders of magnitude and is positively skewed. Applying a logarithmic transform compresses large values and emphasizes smaller ones, producing distributions closer to Gaussian:

$$P_{\log}(f,t) = \log(P(f,t) + \epsilon) \tag{3.7}$$

This transformation is particularly effective in time–frequency analyses, such as event-related spectral perturbations (ERSP) and motor imagery ERD/ERS, where proportional changes are more meaningful than absolute differences [120].

Z-Score Normalization (Standardization)

Z-score normalization (or standardization) removes mean offsets and scales the variance of the EEG signal. For a channel signal $x(t)$ with mean μ and standard deviation σ, the normalized signal becomes:

$$z(t) = \frac{x(t)-\mu}{\sigma} \quad (3.8)$$

This ensures that the signal has zero mean and unit variance, a property particularly useful when EEG features are input to classifiers or regression models. For spectral power features, which are often skewed, a log-transform can be applied before standardization [108].

Min-Max Normalization

Another common transformation rescales the data to a bounded range, often $[0,1]$. If $x(t)$ has minimum and maximum values $min(x)$ and $max(x)$ within the chosen window, the normalized signal is:

$$x'(t) = \frac{x(t)-\min(x)}{\max(x)-\min(x)} \quad (3.9)$$

This approach preserves the shape of the signal but eliminates differences in scale, which is useful for comparing relative amplitude fluctuations across electrodes or trials.

3.2.6.2.6 Data Augmentation

In EEG-based machine learning, especially deep learning, datasets are often limited in size and noisy, which can lead to overfitting. Data augmentation addresses this by artificially increasing the number and diversity of trials while preserving task-related neural information. Augmentation is applied after segmentation, ensuring that the generated trials remain aligned to events such as stimulus onset or motor imagery cues.

Common EEG data augmentation techniques aim to increase the diversity and number of trials while preserving task-related neural information. Time shifting or sliding windows slightly shift epochs in time to create overlapping trials, providing more samples from the same recording. Noise addition introduces small Gaussian noise to simulate measurement variability, while channel perturbation scales or masks electrode signals to mimic variability across electrodes. In the frequency domain, augmentation can involve introducing variations in spectral power or phase to generate alternative but realistic versions of the original signals. Finally, mixup or synthetic trials combine signals from multiple epochs to create new samples, further enhancing dataset diversity. These methods collectively improve model generalization, particularly for deep learning approaches, by increasing dataset size and reducing overfitting [126, 127, 128].

3.2.6.3 Feature Extraction

3.2.6.3.1 Time-Domain Features

Time-domain features are directly extracted from the EEG waveform and capture the temporal characteristics of neural activity. They are simple to compute and provide useful information about the amplitude, variability, and shape of EEG signals during motor imagery tasks. Common examples include mean and variance, which represent the average signal amplitude and its variability within an epoch; root mean square (RMS), which quantifies overall signal power and is sensitive to amplitude changes; Hjorth parameters, including activity (variance of the signal), mobility (square root of the variance of the first derivative divided by activity), and complexity (comparison of the mobility of the first derivative to that of the signal), which, while computed in the time domain, capture aspects of the signal's oscillatory behavior and rate of change; and movement-related potentials (MRPs), which reflect slow cortical potentials associated with motor planning and execution, typically observed as negative or positive

deflections preceding or during motor imagery. Time-domain features are often combined with frequency or spatial features to improve the discriminative power of classifiers and are computationally efficient, making them suitable for real-time BCI systems [53, 129].

3.2.6.3.2 Frequency-Domain Features

Frequency-domain features characterize the oscillatory activity of EEG signals, which is particularly important in motor imagery tasks where rhythmic changes in specific bands are associated with movement intentions. These features are extracted by transforming the time-domain EEG signal into the frequency domain using techniques such as Fourier transform or power spectral density estimation. The most analyzed frequency bands in motor imagery BCI are the Mu (8–13 Hz) and Beta (13–30 Hz) rhythms, which are modulated during imagined or executed movements. Key frequency-domain features include band power, which measures the signal energy within a specific frequency band, and ERD/ERS, which quantifies task-related decreases or increases in band power relative to a baseline period. These features are often averaged over time windows and electrodes or combined with spatial filters to enhance discriminability [53, 129, 130]. A key step in extracting these features is estimating the power spectral density (PSD), which quantifies how signal power is distributed across frequencies. PSD can be computed using segment-averaging methods, such as Welch's method, or direct Fourier-based approaches.

Welch's Method

Welch's method [120, 131] is a widely used approach for PSD estimation that reduces variance while maintaining frequency resolution. The spectral analysis begins with the Discrete Fourier Transform (DFT) of an epoch $x[n]$ of length N,

$$X(f) \;=\; \sum_{n=0}^{N-1} x\,[n]\, e^{-j2\pi f n/N} \tag{3.10}$$

and the (raw) periodogram is the squared magnitude of the DFT (normalized by N),

$$P_{xx}(f) \;=\; \frac{1}{N}\,|X(f)|^2 \tag{3.11}$$

which provides a frequency-wise power estimate but suffers from high variance. To reduce variance, the epoch is divided into L (possibly overlapping) segments of length N_s; denote the k-th segment by $x_k[n]$ for $k = 0, \dots, L-1$. Each segment is multiplied by a window $w[n]$ (e.g., Hann or Hamming) to reduce spectral leakage, producing the windowed segment $w[n]x_k[n]$; the periodogram of the windowed segment (with window-energy normalization) is

$$P_k(f) \;=\; \frac{1}{U}\left|\sum_{n=0}^{N_s-1} w\,[n]\, x_k[n]\, e^{-j2\pi f n/N_s}\right|^2 \tag{3.12}$$

$$U = \frac{1}{N_s}\sum_{n=0}^{N_s-1} w\,[n]^2 \tag{3.13}$$

and Welch's PSD estimate is the average of these segment periodograms,

$$\hat{P}_{xx}(f) \;=\; \frac{1}{L}\sum_{k=0}^{L-1} P_k\,(f) \tag{3.14}$$

where the hat denotes an estimate from finite data. The subscript notation xx identifies this as an auto-spectrum (the spectrum of x with itself); formally the (theoretical) PSD is the Fourier transform of the autocorrelation $R_{xx}(\tau)$, i.e. $P_{xx}(f) = \mathcal{F}R_{xx}(\tau)$, and contrasts with a cross-

spectrum $P_{xy}(f)$ which would describe the spectral relationship between two different signals x and y. Finally, to obtain a single representative value for a given frequency band $[f_{\min}, f_{\max}]$ one may compute either the total band power

$$P_{\text{band}} = \sum_{f=f_{\min}}^{f_{\max}} \hat{P}_{xx}(f)\,\Delta f \tag{3.15}$$

(or the corresponding integral), or the average PSD across the band (commonly reported so values are comparable across bands of different widths),

$$\text{PSD}_{\text{band}} = \frac{1}{f_{\max}-f_{\min}} \sum_{f=f_{\min}}^{f_{\max}} \hat{P}_{xx}(f)\,\Delta f \tag{3.16}$$

where Δf is the frequency-bin width (e.g. $\Delta f = f_s/N_{\text{fft}}$ for sampling frequency f_s and chosen FFT length N_{fft}).

FFT Band-Sum

For short epochs or pre-segmented data (e.g., 0.4 s EEG segments at 160 Hz), it is also common to compute the FFT over each segment or the full epoch and directly sum the squared absolute values of the Fourier coefficients within each frequency band. This produces a single band-power value per electrode per frequency band without further segmentation or Welch averaging:

$$P_{band} = \sum_{k=k_{\min}}^{k_{\max}} |X_e[k]|^2 \tag{3.17}$$

where $X_e(f_k)$ is the FFT of the signal. This approach provides a simple, computationally efficient measure of band power suitable for relative comparisons across electrodes or experimental conditions [46, 120].

3.2.6.3.4 Time-Frequency Domain

While frequency-domain features capture the overall power in specific bands, they do not provide information about when these spectral changes occur. Time–frequency analysis represents EEG signals simultaneously in time and frequency, allowing the detection of transient or event-related oscillatory patterns. This is particularly important in motor imagery, where rhythmic activity in the Mu and Beta bands can appear briefly and vary across trials. Common time–frequency methods include the Short-Time Fourier Transform (STFT) and Discrete Wavelet Transform (DWT).

Short-Time Fourier Transform (STFT)

Motor Imagery (MI) EEG signals are inherently non-stationary, as their spectral properties vary with time depending on cognitive and motor tasks. Classical Fourier Transform assumes signal stationarity and thus fails to capture localized temporal changes in frequency content. To overcome this limitation, the Short-Time Fourier Transform (STFT) is employed, which provides a joint time–frequency representation of the signal and has been extensively used for feature extraction in EEG-based Brain–Computer Interfaces (BCIs) [36, 37].

Given a discrete-time EEG signal $x[n]$, STFT segments the signal into overlapping windows and computes the Fourier transform within each window. The STFT of $x[n]$ is defined as:

$$X(n,\omega) = \sum_{m=-\infty}^{\infty} x[m]\, w[n-m]\, e^{-j\omega m} \tag{3.18}$$

where $w[n]$ is a windowing function (commonly Hamming or Hanning) applied to localize the signal in time, and ω represents the angular frequency. Equation (1) produces a two-dimensional function $X(n, \omega)$, which indicates the frequency content of the signal around time n.

The squared magnitude of the STFT defines the spectrogram,

$$P(n, \omega) = |X(n, \omega)|^2 \tag{3.19}$$

which provides a time–frequency energy distribution and is widely used to extract discriminative patterns from EEG signals. In the context of MI, desynchronization in the mu rhythm (8–13 Hz) and Beta rhythm (13–30 Hz) can be tracked over time using spectrograms, allowing the detection of motor-related activity [46, 52].

In practice, the EEG signal is divided into overlapping segments of length L, each multiplied by the chosen window function. The discrete Fourier transform (DFT) is then applied to each segment, yielding:

$$X_k[n] = \sum_{m=0}^{L-1} x\,[m + nH]\, w[m]\, e^{-j2\pi km/L} \tag{3.20}$$

where H is the hop size (step between adjacent windows), k is the frequency bin index, and n is the frame index [120].

For feature extraction, spectral power within specific frequency bands is typically integrated from the spectrogram. For example, Mu and Beta band powers over time can be extracted as:

$$F_b[n] = \sum_{\omega \in b} P\,(n, \omega) \tag{3.21}$$

where b denotes the frequency band of interest. These band power features are then used as inputs to classifiers such as LDA or SVM.

STFT thus provides a powerful framework for analyzing the non-stationary characteristics of EEG signals. However, it suffers from the time–frequency resolution trade-off, as the choice of window size affects the precision in both domains. Despite this limitation, STFT has been successfully applied in MI-EEG research due to its ability to reveal temporal dynamics in oscillatory brain rhythms [37, 46, 52].

Discrete Wavelet Transform (DWT)

The Discrete Wavelet Transform (DWT) is a powerful tool for analyzing non-stationary EEG signals, as it provides simultaneous time–frequency localization. Unlike the Short-Time Fourier Transform (STFT), which uses a fixed analysis window, the DWT adapts resolution depending on the frequency: fine time resolution at high frequencies and fine frequency resolution at low frequencies. This property makes the DWT particularly suitable for feature extraction in Motor Imagery (MI) EEG, where oscillatory activity varies dynamically across different frequency bands [36, 37].

In DWT, a discrete-time EEG signal $x[n]$ is successively passed through a pair of low-pass and high-pass filters followed by down sampling by a factor of two. This results in a hierarchical decomposition of the signal into approximation and detail coefficients. At decomposition level j, the approximation coefficients $A_j[n]$ and detail coefficients $D_j[n]$ are defined as:

$$A_j[n] = \sum_k h\,[k]\, A_{j-1}[2n - k] \tag{3.22}$$

$$D_j[n] = \sum_k g\,[k]\, A_{j-1}[2n-k] \tag{3.23}$$

where $h[k]$ and $g[k]$ represent the coefficients of the low-pass and high-pass filters, respectively, and $A_0[n] = x[n]$ is the original EEG signal [120].

Through iterative decomposition, the EEG signal is separated into different sub-bands that correspond to physiologically relevant frequency ranges [46]. This decomposition enables the extraction of features directly linked to motor-related brain activity.

From the detail and approximation coefficients, several statistical features can be derived. One of the most widely used is the wavelet energy, which quantifies the signal power in each frequency band:

$$E_j = \sum_n \left|D_j[n]\right|^2 \tag{3.24}$$

where E_j denotes the energy of detail coefficients at level j. Normalization of energy across levels is often performed to ensure scale invariance, defined as:

$$\tilde{E}_j = \frac{E_j}{\sum_{l=1}^{J} E_l} \tag{3.25}$$

where J is the maximum decomposition level. The feature vector is then formed as:

$$f = \left[\tilde{E}_1, \tilde{E}_2, \ldots, \tilde{E}_J\right] \tag{3.26}$$

which captures the distribution of spectral power across wavelet sub-bands.

The choice of the mother wavelet (e.g., Daubechies, Symlets, Coiflets) and the number of decomposition levels significantly influences classification performance. When combined with classifiers such as Linear Discriminant Analysis (LDA) or Support Vector Machines (SVM), DWT features have shown robust performance in EEG-based BCIs [36,37, 92].

3.2.6.5 Spatial Domain

Motor Imagery (MI) EEG signals are often contaminated by noise, artifacts, and overlapping background brain activity, which makes it difficult to extract features that are truly discriminative between different imagined motor tasks. Spatial filtering methods are therefore widely employed to enhance the signal-to-noise ratio (SNR) by linearly combining signals from multiple electrodes in order to emphasize activity from the motor cortex and suppress irrelevant activity. Among these methods, the Common Spatial Patterns (CSP) algorithm has become one of the most widely used and effective approaches in MI-EEG based Brain-Computer Interfaces (BCIs) due to its ability to maximize variance differences between two classes [132].

Let an EEG trial be represented as a matrix $X \in \mathbb{R}^{C\times T}$, where C is the number of channels and T is the number of time samples. CSP seeks a spatial filter matrix $W \in \mathbb{R}^{C\times C}$ such that the transformed signals:

$$Z = W^T X \tag{3.27}$$

maximize variance for one class while minimizing it for the other. To achieve this, the normalized spatial covariance matrix of each trial is computed as:

$$\Sigma = \frac{XX^T}{\text{trace}(XX^T)} \tag{3.28}$$

For a binary classification problem, the mean covariance matrices for class 1 and class 2 are denoted as Σ_1 and Σ_2, respectively, and the CSP algorithm solves the generalized eigenvalue problem:

$$\Sigma_1 w = \lambda \Sigma_2 w \tag{3.29}$$

where the eigenvectors w correspond to spatial filters and the eigenvalues λ reflect the variance ratio between the two classes [132]. The spatial filter matrix W is constructed from a selection of eigenvectors associated with the largest and smallest eigenvalues, thus maximizing variance for one class while minimizing it for the other.

Once the data are spatially filtered, features are extracted by computing the log-variance of the selected components. Specifically, for each component Z_i, the feature is given by:

$$f_i = log\left(\frac{\text{var}(Z_i)}{\sum_j \text{var}(Z_j)}\right) \tag{3.30}$$

which normalizes the variance relative to the total power across components, ensuring scale-invariance [108]. These features are then used as inputs to standard classifiers such as Linear Discriminant Analysis (LDA) or Support Vector Machines (SVM).

Although CSP is highly effective in binary MI-EEG classification, it has several limitations, including sensitivity to noise, session-to-session variability, and the risk of overfitting with small datasets. To overcome these issues, multiple extensions have been proposed. Regularized CSP (RCSP) introduces shrinkage and regularization strategies for covariance estimation to improve robustness [133]. Filter Bank CSP (FBCSP) extends CSP to multiple frequency bands, extracting features from each band and selecting the most discriminative ones [130]. Sparse CSP incorporates sparsity constraints into the optimization to reduce overfitting and improve interpretability [134]. Stationary CSP (sCSP) accounts for non-stationarity across recording sessions to improve generalization [135]. These variants highlight the adaptability of CSP and its continued relevance in modern BCI systems.

In summary, CSP remains a cornerstone technique for feature extraction in MI-EEG due to its discriminative power, computational efficiency, and neurophysiological interpretability. Despite its limitations, ongoing refinements such as RCSP, FBCSP, sparse CSP, and stationary CSP continue to expand its applicability in real-world BCI systems.

3.2.6.4 Classification

Once relevant EEG features are extracted, the next step in a motor imagery BCI pipeline is classification, where patterns corresponding to different mental states are automatically recognized. Traditional classifiers such as Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM) are widely used due to their simplicity, efficiency, and strong performance in BCI applications. Deep Learning approaches, including convolutional and recurrent neural networks, are increasingly applied to EEG classification due to their ability to learn hierarchical representations directly from raw or preprocessed signals, reducing the need for manual feature engineering.

Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is one of the most widely used classifiers in Brain–Computer Interface (BCI) systems due to its simplicity, computational efficiency, and effectiveness in distinguishing Motor Imagery (MI) EEG features [136, 137]. The central idea

of LDA is to find a linear projection that maximizes the separability between different classes by considering both the mean differences and within-class variability of the data.

Let a dataset consist of N samples, each represented as a feature vector $x \in \mathbb{R}^d$, belonging to one of two classes $\mathcal{C}_1$ or $\mathcal{C}_2$. LDA seeks a projection vector $w \in \mathbb{R}^d$ such that the scalar projection:

$$y = w^T x \tag{3.31}$$

maximizes the separation between the projected class means while minimizing the variance within each class. To achieve this, the mean vectors of the two classes are computed as:

$$m_i = \frac{1}{N_i} \sum_{x \in \mathcal{C}_i} x, \quad i \in 1,2 \tag{3.32}$$

where N_i is the number of samples in class $\mathcal{C}_i$. The within-class scatter matrix is defined as:

$$S_w = \sum_{i=1}^{2} \sum_{x \in \mathcal{C}_i} (x - m_i)\,(x - m_i)^T \tag{3.33}$$

and the between-class scatter matrix as:

$$S_b = (m_1 - m_2)(m_1 - m_2)^T \tag{3.34}$$

The optimal projection vector w is then obtained by maximizing the Fisher criterion, defined as the ratio of between-class variance to within-class variance:

$$w = arg\max_{w} \frac{w^T S_b w}{w^T S_w w} \tag{3.35}$$

This optimization problem can be solved as a generalized eigenvalue problem, and in the two-class case, the solution reduces to $w = S_w^{-1}(m_1 - m_2)$. Once the projection vector is found, classification is performed by projecting a test sample x onto w and applying a threshold, typically chosen as the midpoint between the projected class means.

LDA is particularly well-suited for CSP features in MI-EEG, as CSP produces log-variance features that are approximately Gaussian distributed, which aligns with LDA's assumption of normally distributed data with equal covariance across classes [138]. Despite its simplicity, LDA has consistently shown strong performance in BCI competitions and remains a standard benchmark method. However, its assumptions limit its performance when class covariances are significantly different or when data are highly non-linear. Extensions such as Regularized LDA (RLDA) [139], Heteroscedastic LDA (HLDA) [140], and Kernel LDA (KLDA) [141] have been proposed to address these limitations by improving robustness, handling unequal covariance structures, and capturing non-linear class boundaries, respectively. LDA provides an elegant balance of mathematical simplicity, computational efficiency, and classification accuracy, which explains its long-standing role as a baseline classifier in MI-EEG based BCIs.

Support Vector Machine (SVM)

Support Vector Machines can be broadly categorized based on the type of decision boundary and kernel function used. The linear SVM is suitable when classes are linearly separable and computational efficiency is a priority [142, 143]. For data that exhibits non-linear relationships, non-linear SVMs employ kernel functions to implicitly map inputs into higher-dimensional feature spaces. Common kernels include polynomial, sigmoid, and Radial Basis Function (RBF).

The RBF kernel is often preferred due to its ability to handle highly complex and non-linear decision boundaries with only a single additional hyperparameter, γ. It measures similarity based on the Euclidean distance between points, creating a localized influence around each support vector [144]. This makes RBF-SVM highly flexible, robust to high-dimensional features, and effective in applications such as brain–computer interface (BCI) classification, where the underlying patterns in features like Power Spectral Density (PSD) are rarely linear. Given its balance of flexibility, generalization, and computational tractability, the RBF-SVM is widely adopted for tasks where class boundaries are intricate and the data is high-dimensional, motivating its use in this work. Support Vector Machines are particularly attractive for classification tasks because they are grounded in convex optimization, which guarantees a unique global solution, and because the resulting classifier is defined only by a subset of the training points, the *support vectors*. This sparsity improves both efficiency and generalization, which are critical in applications such as brain–computer interfaces (BCIs), where high-dimensional features must be processed with limited data.

The soft-margin SVM is usually formulated in its dual form, where the optimization problem depends only on the inner products of training samples [145]. The dual problem is expressed as:

$$\max_{\alpha} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \, \alpha_j y_i y_j K(x_i, x_j) \tag{3.36}$$

subject to $0 \leq \alpha_i \leq C \quad$ for $i = 1, \ldots, n \quad$ and $\sum_{i=1}^{n} \alpha_i \, y_i = 0$. Here, the parameter $C > 0$ controls the trade-off between maximizing the margin and minimizing classification errors. A large C penalizes misclassifications more heavily, leading to a narrower margin and potential overfitting, while a smaller C allows for a wider margin but may tolerate some misclassified points.

To handle non-linear separability, the dot product in the dual problem is replaced with a kernel function $K(x_i, x_j)$. In practice, the *Radial Basis Function (RBF) kernel* has proven particularly effective because it can model highly non-linear class boundaries with only one additional parameter γ. It is defined as:

$$K(x_i, x_j) = exp(-\gamma \parallel x_i - x_j \parallel^2) \tag{3.37}$$

where γ determines the locality of influence of each support vector. A small γ produces a smoother decision boundary by considering distant points as similar, while a large γ makes the model highly sensitive to individual points, risking overfitting.

Once the optimization problem is solved, the classifier is expressed in terms of the support vectors and their corresponding multipliers. The resulting decision function for a new sample x is:

$$f(x) = \sum_{i=1}^{n} \alpha_i \, y_i exp(-\gamma \parallel x_i - x_j \parallel^2) + b \tag{3.38}$$

where only points with nonzero α_i contribute to the classification. The bias term b can be recovered using any support vector x_k satisfying $0 < \alpha_k < C$:

$$b = y_k - \sum_{i=1}^{n} \alpha_i \, y_i exp(-\gamma \parallel x_i - x_j \parallel^2) \tag{3.39}$$

Finally, the predicted class label is determined by the sign of the decision function:

$$\hat{y} = \text{sign}\big(f(x)\big) \tag{3.40}$$

In summary, the performance of RBF-SVM hinges on the careful selection of its two hyperparameters: C, which balances margin width and misclassification, and γ, which controls the shape of the decision boundary. Together, they determine whether the model underfits, overfits, or generalizes well. This combination of theoretical robustness, non-linear flexibility, and tunable hyperparameters makes the RBF-SVM a highly feasible and effective classifier, particularly in domains such as signal processing and BCI, where decision boundaries are rarely linear and robustness is essential.

Chapter 4

Signal to Image Transformation and Deep Learning

4.1 Introduction

Despite major advances in brain–computer interface (BCI) research, motor imagery (MI) EEG classification remains a difficult challenge. EEG signals are low in signal-to-noise ratio, highly non-stationary, and subject to strong inter-trial and inter-subject variability. Data collection is also resource-intensive, leading to relatively small datasets that complicate robust feature extraction. Hence, the choice of classification strategy is crucial. While traditional methods such as adaptive classifiers [146], Riemannian geometry [147,148], and tensor-based approaches have demonstrated strong performance, deep learning [149, 150] has recently attracted increasing attention due to its ability to automatically extract hierarchical spatial–spectral–temporal features from raw or minimally processed EEG.

One promising direction is *signal-to-image transformation*, where EEG is mapped into structured 2D or 3D representations that preserve spatiotemporal and spectral dynamics. This allows deep networks such as CNNs, LSTMs, and more recently transformers, to operate on EEG in ways analogous to computer vision. For example, Qiao and Bi [151] combined an inception-based CNN with a bidirectional GRU to classify four MI tasks from the BCI Competition IV 2a dataset, using Morlet wavelets with cubic spline interpolation. Xu et al. [152] applied a two-layer CNN on wavelet-based images from C3/C4 electrodes, while Chaudhary et al. [153] used transfer learning with pre-trained CNNs on EEG scalograms to overcome data scarcity. Ha and Jeong [16] demonstrated that capsule networks operating on STFT-based images achieved higher accuracy for binary MI classification.

A wide range of transformation techniques exist. Classical time–frequency methods include the short-time Fourier transform (STFT), which produces spectrograms but suffers from fixed resolution, and the continuous wavelet transform (CWT), which offers multi-resolution analysis tailored to EEG's non-stationarity. Beyond these, novel encodings have emerged: Gramian Angular Fields (GAF) and Markov Transition Fields (MTF) preserve temporal correlations and state transitions [154, 155]; Recurrence Plots (RP) visualize nonlinear phase-space recurrences, recently combined with graph neural networks [156]; and polar coordinate embeddings such as Angle-Amplitude Transformations (AAT) emphasize amplitude–phase dynamics [157]. More recently, vision-transformer architectures[158], ST-ViT [159] treat EEG topographic maps as patch embeddings, applying self-attention to capture long-range inter-electrode dependencies.

High-density montages (64–256 electrodes, e.g., BioSemi ActiveTwo, Neuroscan SynAmps II, EGI HydroCel) allow for accurate spatial mapping via interpolation, beamforming, or source localization (LORETA, sLORETA). Importantly, spatial filtering methods such as CSP, FB-CSP, and RCSP [130, 160, 161] remain complementary. Recent work such as CSP-Net [162] integrates CSP layers within CNN architectures—either by placing a CSP layer before a CNN

backbone or by replacing a convolutional layer with a learnable CSP filter — showing improved performance particularly in small-sample motor imagery classification settings.

Emerging trends further push the boundaries. Domain adaptation techniques, including adversarial learning [163], subject-specific normalization [164], and few-shot meta-learning [165], improve cross-subject generalization—an essential step for clinical translation. Multimodal fusion approaches integrate EEG with MEG and leverage Attention Neural Networks for improved brain source imaging [166], while graph neural networks [167] capture dynamic brain connectivity. Data augmentation via diffusion models [168] promises to alleviate dataset scarcity by generating realistic synthetic EEG.

Most existing studies simplify motor imagery classification by using only a small subset of electrodes or restricted datasets, which limits the ability to generalize across subjects and sessions. In contrast, our approach leverages the richness of a high-density 64-channel montage, incorporating the entire electrode array into the signal-to-image transformation process. Instead of focusing narrowly on a few contralateral sites, the full-montage projection captures distributed cortical activity, including subtle inter-hemispheric interactions and secondary motor areas often overlooked in reduced setups.

To capture spectral diversity, we decompose EEG into Mu (8–12 Hz) and Beta (18–30 Hz) rhythms, the canonical sensorimotor bands associated with event-related desynchronization/synchronization during motor imagery. In addition, we include the Delta band (0.5–4 Hz), which is traditionally considered irrelevant in MI paradigms due to its association with slow cortical processes and artifacts. However, by explicitly incorporating Delta activity as a separate image channel, we investigate whether it carries complementary information—such as baseline cortical state or low-frequency modulations—that can enhance class separability. Demonstrating discriminative potential in the Delta band would challenge conventional assumptions and expand the spectral scope of MI-based BCI research.

To preserve spectral diversity, we decompose the EEG into multiple physiologically relevant bands—Delta, mu, and Beta—and encode them as complementary image channels, analogous to RGB layers in natural images. This multi-band representation not only highlights the canonical Mu and Beta rhythms linked to motor preparation and execution, but also incorporates low-frequency Delta dynamics that may reflect baseline modulation and global cortical state.

These images are generated as topographic maps, interpolated across the scalp to reflect continuous spatial patterns rather than isolated electrode signals. The resulting representations retain both fine-grained local activations (e.g., contralateral ERD in C3/C4) and global network-level organization.

Deep learning models then operate on these structured images. Convolutional neural networks (CNNs) hierarchically extract spatial features, identifying discriminative patterns such as localized ERD/ERS or hemispheric asymmetries, while long short-term memory networks (LSTMs) capture temporal dependencies across MI trials, modeling how oscillatory dynamics evolve over time. By combining these architectures, the framework jointly exploits spatial, spectral, and temporal cues for robust five-class MI classification (left fist, right fist, both fists, both feet, rest).

4.2 Deep Learning Models

4.2.1 Artificial Neural Network (ANN)

Artificial Neural Networks (ANNs) are computational models inspired by the biological brain, designed to approximate complex functions. At their core, they consist of layers of interconnected nodes (neurons). Each neuron performs a simple transformation: it takes inputs, applies weights and biases, and passes the result through a nonlinear activation function [169].

For a single layer, the computation can be written as:

$$z = Wx + b, \quad h = \phi(z) \tag{4.1}$$

where: $x \in R^{d_{\text{in}}}$: input vector, $W \in R^{d_{out} \times d_{\text{in}}}$: weight matrix, $b \in R^{d_{out}}$: bias vector, z: pre-activation output, $\phi(\cdot)$: nonlinear activation (e.g., ReLU, sigmoid, tanh), $h \in R^{d_{out}}$: neuron outputs (activations).

By stacking multiple such layers, neural networks can learn hierarchical feature representations: earlier layers capture low-level features, while deeper layers capture abstract patterns.

Neural networks learn by adjusting weights W and biases b to minimize a discrepancy between predicted outputs and true labels.

Let training data be:

$$\mathcal{D} = \{(\mathrm{x}^{(i)}, \mathrm{y}^{(i)})\}_{i=1}^{N} \tag{4.2}$$

where $\mathrm{x}^{(i)}$ is the input and $\mathrm{y}^{(i)}$ is the target output.

The network produces predictions $\hat{\mathrm{y}}^{(i)} = f_\theta(\mathrm{x}^{(i)})$, where θ denotes all parameters (weights and biases).

The loss function quantifies the error between predictions and targets. Common examples:

Mean Squared Error (MSE) for regression:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N} \| \mathrm{y}^{(i)} - \hat{\mathrm{y}}^{(i)} \|^2 \tag{4.3}$$

Cross-Entropy Loss for classification with C classes:

$$\mathcal{L}_{\text{CE}} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_c^{(i)} \log\left(\hat{y}_c^{(i)}\right) \tag{4.4}$$

where $y_c^{(i)}$ is 1 if sample i belongs to class c, otherwise 0.

Learning consists of iteratively minimizing the loss function with respect to network parameters. This is typically achieved through Stochastic Gradient Descent (SGD) and its variants.

The gradient of the loss with respect to parameters is computed via backpropagation:

$$\nabla_\theta \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \theta} \tag{4.5}$$

Parameters are updated as:

$$\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L} \tag{4.6}$$

where: θ: all trainable weights and biases, η: learning rate (step size), $\nabla_\theta \mathcal{L}$: gradient of loss with respect to parameters.

Variants such as Adam, RMSProp, and momentum-based SGD improve convergence by adapting the learning rate or incorporating past gradients.

4.2.2 Convolutional Neural Network (CNN)

The concept of convolutional neural networks (CNNs) originated from the biological inspiration of the visual cortex, particularly the discovery of receptive fields in the work of Hubel and Wiesel [170]. Early computational models, such as Neocognitron [171], introduced the idea of local connectivity and weight sharing to mimic hierarchical feature extraction. Modern CNNs were popularized by LeCun et al. [172] through the LeNet architecture, which successfully recognized handwritten digits. Later, the breakthrough AlexNet [173] demonstrated the effectiveness of deep CNNs for large-scale image classification, significantly outperforming traditional handcrafted feature approaches. Building on these advances, deeper architectures such as VGGNet [174] and ResNet [175] refined convolutional designs and demonstrated that depth, with careful architectural choices, could yield superior generalization. CNNs have since become the dominant architecture for computer vision and related tasks, including medical imaging, speech processing, and EEG-based classification. Their strength lies in automatic feature extraction through hierarchical filters, enabling models to learn low-level features (edges, textures) and high-level representations (objects, patterns) directly from raw data. A CNN learns hierarchical representations through a series of learnable filters and non-linear transformations. The core components include convolutional layers, activation functions, pooling operations, and fully connected layers.

1. Convolutional Layer

Given an input image (or feature map) $X \in R^{H \times W \times C_{\text{in}}}$, where H is the height, W is the width, and C_{in} is the number of input channels, a convolutional layer applies a set of learnable kernels (filters) $R^{k \times k \times C_{\text{in}} \times C_{\text{out}}}$ where k is the kernel size and C_{out} is the number of output channels.

The convolution operation at spatial location (i, j) for output channel mmm is:

$$Y_{i,j,m} = \sum_{c=1}^{C_{\text{in}}} \sum_{u=1}^{k} \sum_{v=1}^{k} K_{u,v,c,m} \cdot X_{i+u,j+v,c} + b_m \tag{4.7}$$

where b_m is the bias term for channel m. Stride s: controls the step size of the convolution. Padding p: adds zeros around the input to preserve spatial resolution

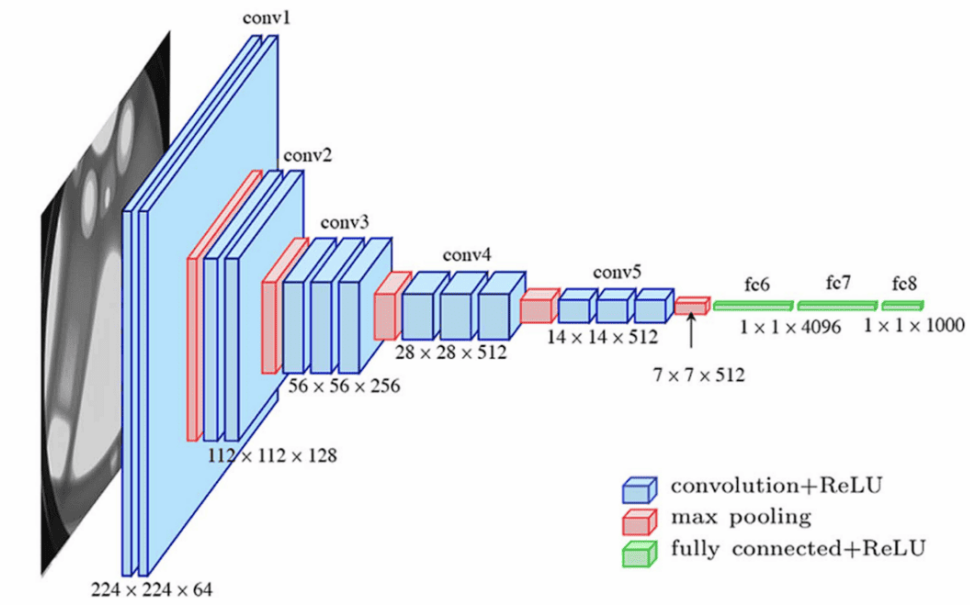

Figure 8: Architecture of the VGG Convolutional Neural Network. Schematic representation of the VGG network illustrating the sequence of convolutional layers, pooling layers, and fully connected layers. The model is characterized by its use of small 3×3 convolutional filters, stacked in increasing depth across the network, with max-pooling layers for spatial downsampling and fully connected layers for classification at the output [174].

2. Activation Function

After convolution, a non-linear activation is applied. The most common is the Rectified Linear Unit (ReLU):

$$\phi(z) = ma\,x(0, z)\,.\phi(z) \tag{4.8}$$

ReLU introduces non-linearity while being computationally efficient and alleviating the vanishing gradient problem compared to sigmoid or tanh activations.

3. Pooling Layer

Pooling reduces the spatial resolution, enabling translation invariance and reducing computational complexity. The most common is MaxPooling, defined as:

$$Y_{i,j,c} = \max_{(u,v)\in\mathcal{R}} X_{s\cdot i+u,\, s\cdot j+v,\, c} \tag{4.9}$$

where R is the pooling region (e.g., 2 × 2), and s is the stride.

4. Fully Connected Layer

A fully connected (dense) layer flattens the feature maps and applies a linear transformation:

$$y = Wx + by = Wx + b \tag{4.10}$$

where:

$$W \in R^{d_{\text{out}} \times d_{\text{in}}}, x \in R^{d_{\text{in}}}, y \in R^{d_{\text{out}}}$$

The final layer for classification applies the softmax function to output probabilities over C classes:

$$\widehat{y_c} = \frac{\exp(z_c)}{\sum_{j=1}^{C} \exp(z_j)}, \quad c = 1, \ldots, C. \tag{4.11}$$

5. Training

CNN parameters are optimized by minimizing a loss function using gradient descent. For classification, the cross-entropy loss is standard:

$$\mathcal{L} = -\sum_{c=1}^{C} y_c \log(\widehat{y_c}) \tag{4.12}$$

where y_c is the one-hot ground truth label and $\widehat{y_c}$ is the predicted probability.

Optimization is performed using Stochastic Gradient Descent (SGD):

$$\theta \leftarrow \theta - \eta \nabla_{\theta \mathcal{L}} \tag{4.13}$$

where θ are the learnable parameters, η is the learning rate, and $\nabla_{\theta \mathcal{L}}$ is the gradient computed by backpropagation.

4.2.3 Long Short-Term Memory (LSTM) Networks

Recurrent Neural Networks (RNNs) were originally designed to process sequential data by maintaining a hidden state that evolves over time. However, traditional RNNs suffer from the vanishing and exploding gradient problem [176], which hinders their ability to capture long-term dependencies. To address this limitation, Sepp Hochreiter and Jürgen Schmidhuber [177] introduced the Long Short-Term Memory (LSTM) architecture. LSTMs augment the RNN structure with a memory cell and gating mechanisms that regulate information flow, enabling stable learning over long sequences. Today, LSTMs are widely used in natural language processing, time series forecasting, and EEG/biomedical signal analysis.

Recurrent Neural Networks (RNNs) are designed to model sequential and time-dependent data, addressing a key limitation of feed-forward neural networks. In standard feed-forward architectures, the prediction $\mathrm{p}(\mathrm{y_t} \mid \mathrm{x_t})$ depends only on the current input $\mathrm{x_t}$, with all prior information encoded implicitly in the learned weights after training. This lack of explicit memory makes it difficult to capture sequential or temporal dependencies.

RNNs extend this paradigm by introducing a hidden state h_t that recurs across time:

$$h_t = f(W_x x_t + W_h h_{t-1} + b) \tag{4.14}$$

where: x_t is the input at time t, h_t is the hidden state, W_x, W_h are weight matrices, b is a bias term, $f(\cdot)$ is a nonlinear activation (e.g., $tanh$ or $ReLU$).

The unfolded representation of an RNN shows how the hidden state is passed along a temporal chain, enabling the network to condition predictions on both current and past information:

$$p(y_t \mid x_t, h_{t-1}) \tag{4.15}$$

This recurrence allows RNNs to retain information from earlier time steps, making them suitable for sequence modeling in language, speech, and EEG signals.

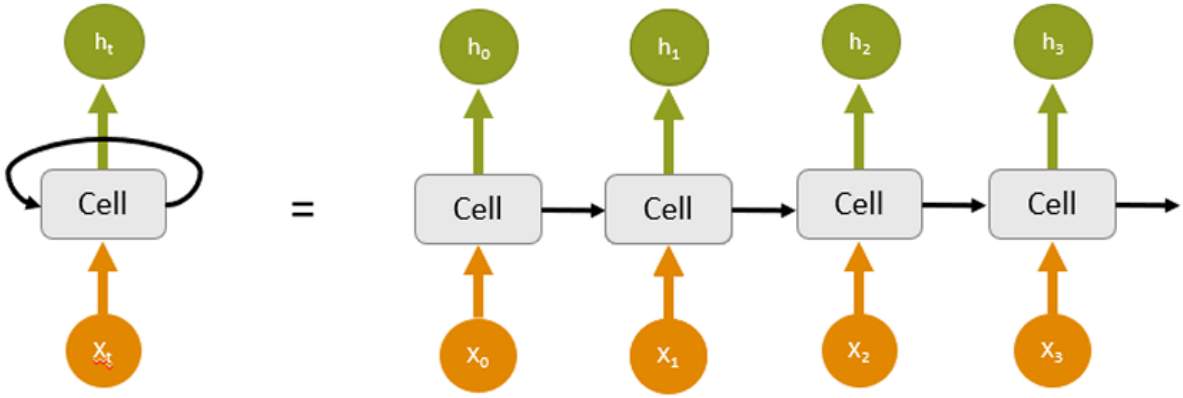

Figure 9: Unfolded Representation of a Recurrent Neural Network (RNN). Schematic illustration of an RNN unfolded through time, showing how the hidden state is passed recurrently from one time step to the next. Each unit receives input at time t, updates its hidden state based on both the new input and the previous hidden state, and produces an output. This unfolded view highlights the temporal dependencies modeled by the network.

The vanishing gradient problem is a major challenge in training RNNs**.** In theory, the hidden state h_t can carry information across an arbitrary number of time steps, enabling long-term dependencies to be modeled. In practice, however, training relies on *Backpropagation Through Time (BPTT)*, where loss gradients must be propagated not only across the depth of the network but also through the temporal dimension. The recursive update $h_t = f(W_x x_t + W_h h_{t-1} + b)$ causes gradients to involve repeated multiplication by Jacobian matrices. Over many time steps, these multiplications tend to either: Vanish (gradients approach 0) leading to long-term dependencies cannot be learned, or Explode (gradients grow uncontrollably) leading to unstable training. This is known as the *vanishing/exploding gradient problem* [176] It is especially problematic for tasks where relevant dependencies span tens or hundreds of time steps, such as EEG signal analysis.

To overcome the vanishing gradient problem, Hochreiter and Schmidhuber [177] proposed the *Long Short-Term Memory (LSTM)* architecture. An LSTM cell enhances the RNN cell by introducing a *cell state* c_t, which serves as long-term memory. Unlike the hidden state, which is updated at every time step, the cell state allows information to flow along a conveyor belt-like pathway, with only minor linear interactions. Thus, the LSTM has access to both *short-term memory*, represented by the hidden state h_t, and *long-term memory*, maintained through the cell state c_t. Carefully designed *gates* (forget, input, output) regulate how information enters, persists, or leaves the cell. These gates ensure that essential information and error gradients can pass unchanged across long temporal distances, solving the vanishing gradient problem.

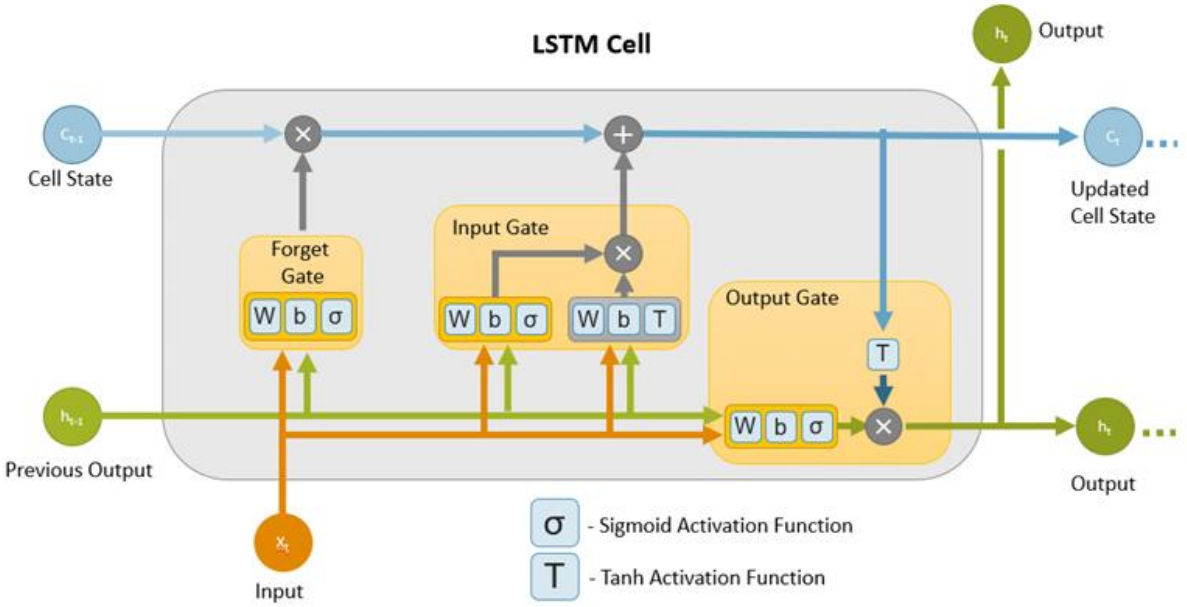

Figure 10: Long Short-Term Memory (LSTM) Cell Structure. Schematic diagram of a single LSTM cell showing the flow of information through the cell state and hidden state. The input gate, forget gate, and output gate regulate how new information is added, how past information is retained or discarded, and how the final hidden state is produced. This gating mechanism enables LSTMs to capture both short- and long-term dependencies in sequential data.

Consider an input sequence: $x_1, x_2, \dots, x_T$, where $x_t \in R^{d_x}$ with hidden state dimension d_h.

At each time step t, the LSTM maintains:

- A hidden state: $h_t \in R^{d_h}$
- A cell state (memory): $c_t \in R^{d_h}$

The dynamics are governed by gates, each represented as a function of the current input x_t ,and previous hidden state h_{t-1}.

1. Forget Gate: Controls how much of the previous memory c_{t-1} should be retained.

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{4.16}$$

where: $f_t \in [0,1]^{d_h}$ (elementwise values close to 0 → forget, close to 1 → keep), $W_f \in R^{d_h \times d_x}$, $U_f \in R^{d_h \times d_h}$, $b_f \in R^{d_h}$.

2. Input Gate: Regulates how much new information enters the celland is defined as:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{4.17}$$

a candidate cell update is then computed as:

$$\tilde{c_t} = \tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{4.18}$$

where $\tilde{c_t} \in R^{d_h}$ represents new candidate memory content.

3.Cell State Update: The new cell state combines retained old memory and candidate information:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{4.19}$$

where $\odot$ denotes elementwise multiplication.

4. Output Gate: Determines which parts of the memory are exposed to the hidden state:

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{4.20}$$

The hidden state is then updated as:

$$h_t = o_t \odot \tanh(c_t) \tag{4.21}$$

Here, $\sigma(\cdot)$ denotes the sigmoid activation function, $\sigma(z) = \frac{1}{1+e^{-z}}$, which squashes values into $(0,1)$ and is used in the gates. The function $tanh(\cdot)$ denotes the hyperbolic tangent activation, $\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$, which outputs values in $(-1,1)$ and is used to regulate candidate updates and hidden states.

4.3 Azimuthal Equidistant Projection for EEG Electrode Layouts

To visualize and analyze scalp EEG data, electrode positions obtained from the realistic head model were projected from three-dimensional head coordinates onto a two-dimensional plane using an azimuthal equidistant projection. This method is well suited for EEG topography because it preserves the relative distances of electrodes from a central reference point. In this study, the vertex electrode (Cz) was selected as the projection center, as it lies near the geometric midpoint of the electrode array and minimizes distortion across the scalp surface. Each electrode position (x, y, z) was first expressed in spherical coordinates, where the radius is defined as $r = \sqrt{x^2 + y^2 + z^2}$, the colatitude as $\arccos(z/r)$, and the azimuth as $\phi = \arctan 2\,(y, x)$. The azimuthal projection then maps these spherical coordinates to two-dimensional Cartesian coordinates through $\rho = k\theta, X = \rho\cos(\phi)$, and $Y = \rho\sin(\phi)$, where ρ represents the distance from Cz in the 2D plane and k is a scaling factor proportional to the head radius. This formulation ensures that distances from Cz are preserved, thereby maintaining local spatial relationships between electrodes. By applying this projection, the 64-electrode configuration derived from the realistic head model was transformed into a two-dimensional representation suitable for generating interpretable electrode layouts and scalp potential maps [178]. The resulting projection is illustrated in Figure. 11.

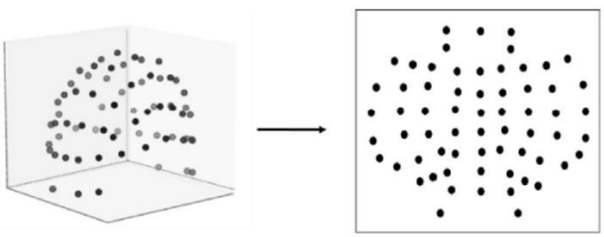

Figure 11: Electrode positions over the realistic head model (left) and their azimuthal equidistant projection onto a 2D plane using Cz as the center (right).

4.4 Interpolation Methods for Scattered Spatial Data

Interpolation is a fundamental technique for estimating values at unsampled locations based on known data points. In EEG and other spatially distributed measurements, electrode positions are typically irregularly spaced, requiring interpolation to generate continuous maps. Several commonly used interpolation methods exist, including nearest-neighbor interpolation, which assigns the value of the closest known point; linear or barycentric interpolation, which constructs piecewise linear surfaces; polynomial and spline-based interpolation, which fit continuous functions with varying degrees of smoothness; and radial basis function (RBF) or Kriging interpolation, which generate smooth surfaces based on distance or spatial correlation [179–181]. While these methods each have advantages, they also have limitations such as blocky discontinuities, non-smooth gradients, over-smoothing, or computational complexity. One approach that addresses many of these challenges is Clough–Tocher interpolation, a piecewise cubic method on triangular meshes that naturally adapts to the convex shape of the electrode array [179]. The following subsections provide a brief overview of nearest-neighbor and linear interpolation, followed by a detailed description of Clough–Tocher interpolation.

1. Nearest-Neighbor Interpolation

Nearest-neighbor interpolation is the simplest method, assigning the value of the closest known data point to each unsampled location. Mathematically, given a set of known data points $\{(x_i, y_i, f_i)\}_{i=1}^{N}, where\ (x_i, y_i)$ are the coordinates and f_i the corresponding values, the interpolated value at a query point (x, y) is defined as:

$$f(x,y) = f_j \quad \text{where} \quad j = \arg\min_i \sqrt{(x-x_i)^2 + (y-y_i)^2} \tag{4.22}$$

In other words, $f(x, y)$ is assigned the value of the closest known point. While computationally efficient and easy to implement, this method produces a blocky, discontinuous surface, making it unsuitable for applications that require smooth topographies, such as EEG scalp maps. Nevertheless, it can serve as a quick approximation or for categorical data where smoothness is not essential [179].

2. Linear Interpolation

Linear interpolation constructs a piecewise linear surface by connecting known data points, often using triangular meshes. Within each triangle defined by vertices $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ with values f_1, f_2, f_3, the interpolated value at a query point (x, y) can be expressed using barycentric coordinates $(\lambda_1, \lambda_2, \lambda_3)$ as:

$$f(x,y) = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 \tag{4.23}$$

where the barycentric coordinates are given by:

$$\lambda_1 = \frac{A\big((x,y),(x_2,y_2),(x_3,y_3)\big)}{A\big((x_1,y_1),(x_2,y_2),(x_3,y_3)\big)} \tag{4.24}$$

$$\lambda_2 = \frac{A\big((x,y),(x_3,y_3),(x_1,y_1)\big)}{A\big((x_1,y_1),(x_2,y_2),(x_3,y_3)\big)} \tag{4.25}$$

$$\lambda_3 = 1 - \lambda_1 - \lambda_2 \tag{4.26}$$

Here, $A\left((x_i, y_i), (x_j, y_j), (x_k, y_k)\right)$ denotes the area of the triangle formed by the three points. This method preserves convexity and is relatively fast, but the resulting surface is only C^0-continuous, meaning that the function is continuous while its gradients are not, which can produce sharp edges in visualizations of spatially continuous data [179, 180].

3.Clough–Tocher Interpolation

Clough–Tocher interpolation is a piecewise cubic interpolation method designed to construct smooth surfaces over scattered data points, particularly on triangular meshes. It is widely used in applications such as EEG scalp topography due to its ability to generate continuous and smooth surfaces from irregularly spaced data. The method begins by performing a Delaunay triangulation of the input points, dividing the domain into non-overlapping triangles. Each triangle is then subdivided into three sub-triangles by connecting the centroid of the triangle to its three vertices. Over each sub-triangle, a cubic polynomial $P(x, y)$ is defined as:

$$P(x,y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 + a_6x^3 + a_7x^2y + a_8xy^2 + a_9y^3 \qquad (4.27)$$

where the coefficients are determined to match the data values at the vertices while ensuring that the first derivatives $\frac{\partial P}{\partial x}$ and $\frac{\partial P}{\partial y}$ are continuous across triangle edges, producing a C^1-continuous surface. The derivatives at the centroids are also chosen to optimize smoothness across the sub-triangles. This approach produces a mathematically smooth surface that avoids artificial sharp edges or discontinuities, making it particularly suitable for spatially continuous data such as EEG measurements. Moreover, because it is defined on triangular meshes, Clough–Tocher interpolation naturally adapts to the convex shape of the electrode array, ensuring that no artificial extrapolation is introduced outside the scalp boundary. When applied to EEG electrode layouts, this method allows discrete electrode measurements to be interpolated into smooth two-dimensional topographic maps that preserve local spatial features, facilitating visualization and further analysis [179].

4.3 Materials and Methods

4.3.1 PhysioNet Motor Movement/Imagery Dataset

We utilized the PhysioNet EEG Motor Movement/Imagery Dataset [91], which contains recordings from 109 subjects collected with the BCI2000 system. EEG signals were acquired from 64 electrodes following the 10–10 system, sampled at 160 Hz.

Each subject performed 14 runs:

- 2 baseline runs (one with eyes open and one with eyes closed), each lasting one minute.
- 12 task runs, corresponding to 4 different motor tasks repeated three times, each run lasting two minutes.

The four task types were:

- Real movement: opening and closing the left or right fist depending on the target location (left or right side of the screen).
- Imagined movement: imagining opening and closing the left or right fist.
- Real movement: opening and closing both fists or both feet.
- Imagined movement: imagining opening and closing both fists or both feet.

Each task run contained 15 trials, each lasting approximately 8 seconds and consisting of two distinct epochs: a nearly 4-second Motor Imagery (MI) epoch, followed immediately by a nearly 4-second Rest epoch. Trials were annotated as:

- T0: Rest epoch (no movement or imagery).
- T1: Motor imagery or real epoch corresponding to either left fist or both fists, depending on the run condition.
- T2: Motor imagery or real epoch corresponding to either right fist or both feet.

After every T1 or T2 epoch, a rest epoch (T0) followed.

Our study focuses on classifying five distinct classes derived from this dataset: imagined left fist movement, imagined right fist movement, imagined both fists movement, imagined both feet movement, and Rest.

Following data quality assessment, subjects S043, S088, S089, S092, S100, and S104 were excluded [182] leaving 103 subjects for the analysis.

This setup enables a comprehensive evaluation of the PhysioNet dataset in terms of MI classification.

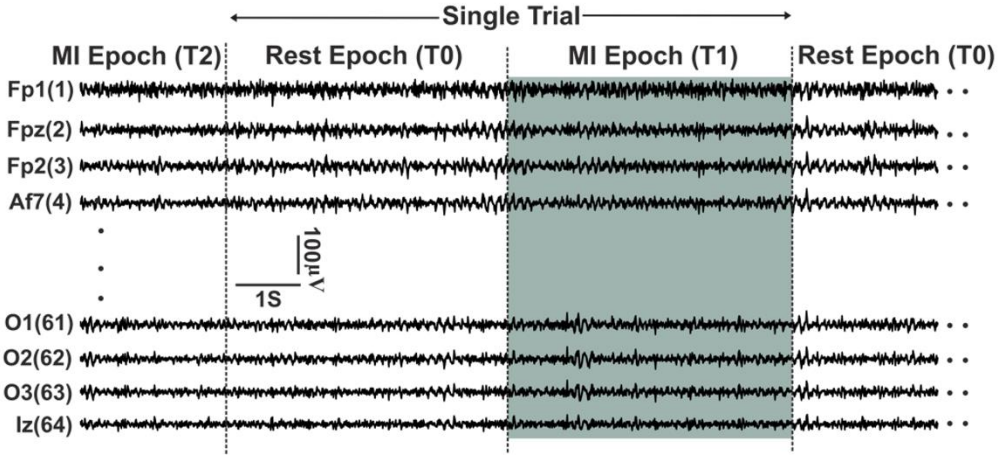

Figure 12: EEG recordings from the PhysioNet Motor Imagery dataset. The vertical axis corresponds to the 64 electrodes arranged according to the international 10–10 system, while the horizontal axis represents time. An illustrative 8-second motor imagery (MI) single trial is shown, consisting of a T0 rest epoch followed by a highlighted T1 MI epoch.

4.3.2 Motor Imagery EEG Signal to Image Transformation

4.3.2.1 Preprocessing

The EEG data were bandpass filtered using a zero-phase 5th-order Butterworth filter into three frequency bands: Delta (0.5–4 Hz), Mu (8–13 Hz), and Beta (13–30 Hz).

Each trial performed by a subject consisted of two epochs: a rest epoch and a motor imagery epoch, each approximately 4 seconds long. Each epoch was further segmented into 10 segments of 0.4 seconds each.

The power spectral density (PSD) for each electrode was computed over every segment using one of two methods, depending on the projection type:

1- the Fast Fourier Transform (FFT), where the squared absolute values of the Fourier coefficients within each frequency band were summed to produce a single band-power value per electrode per frequency band:

$$P_{\text{band}} = \sum_{k=k_{\min}}^{k_{\max}} |X_e[k]|^2 \quad (4.28)$$

where $X_e(f_k)$ is the FFT of the signal at electrode e.

2- Welch's method with a Hann window of 0.4 seconds, where PSD coefficients were averaged within each frequency band to yield the band-power estimate.

4.3.2.1 Chessboard Projection

We propose the Chessboard EEG signal-to-image transformation to convert motor imagery EEG signals into images. The 3D electrode positions are projected onto a 2D chessboard-like grid, where each electrode occupies a 4 × 4 pixel square. This arrangement preserves the relative spatial organization of the electrodes in a simplified 2D grid, making it suitable for image-based analysis. The color intensity of each square encodes the band-power of the corresponding electrode: green shades for Mu band activity and red shades for Beta band activity, allowing quantitative visualization of neuronal activity. Band-power was estimated using the average FFT method, which outperformed Welch's in the pilot phase (described later). This may arise from each electrode mapping independently to a spatial pixel. Sharp spectral contrasts between electrodes likely serve as discriminative features for classifications, which Welch's smoothing could suppress and thus reduce spatial discriminability.

Each unicolor chessboard has a size of 32 × 32 pixels, representing the activity of 64 electrodes for a single frequency band (Figure 13 illustrates the chessboard layout and EEG signal-to-image transformation pipeline). This resolution provides sufficient spatial fidelity for a 64-channel montage while avoiding the redundancy and interpolation artifacts of larger grids. It also aligns with well-established deep learning benchmarks such as CIFAR-10, and as discussed in the following section, it was a key factor in keeping both the hyperparameter screening computationally tractable while enabling full LOOCV evaluation on 103 subjects (described in the next section) and captures subject-specific MI patterns through 10-segment-per-epoch resolution.

The FC5 followed by FC3 ordering ensures primary motor imagery electrodes—C3, Cz, and C4 (sensorimotor cortex, strongest ERD/ERS)—sit in spatially coherent grid positions, letting CNN local filters capture their joint patterns within single receptive fields.

By combining the Mu and Beta chessboards, 2-channel images are obtained. For each subject, the total number of images was calculated as 30 tasks (epochs) × 10 segments per task × 6 motor imagery runs, resulting in 1,800 images per subject. Across all 103 subjects, this amounted to 1,800 × 103 = 185,400 images in total.

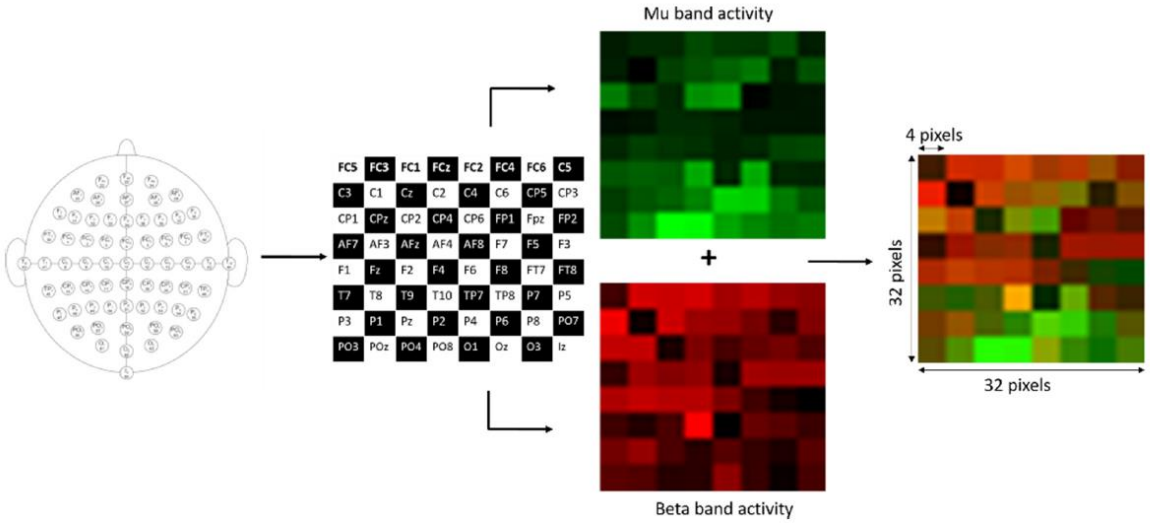

Figure 13: Chessboard EEG signal-to-image transformation method. The 64 electrodes are projected onto a 2D chessboard-like grid, where each electrode occupies a 4 × 4 pixel square, forming a 32 × 32 2-channel image. Mu band activity is represented in green shades and Beta band activity in red shades, with the color intensity of each square reflecting the band-power of the corresponding electrode.

Deep Learning Model

For classifying the Chessboard projection images, we designed a hybrid CNN–LSTM network. The CNN branch, inspired by VGGNet [174], extracts spatial features, while the LSTM branch [177] captures temporal dependencies in the EEG sequences.

Cross-validation is a widely used model evaluation technique in which the available dataset is partitioned into disjoint subsets for training, validation, and testing, allowing performance to be assessed on unseen data while reducing overfitting. Among these approaches, Leave-One-Out Cross-Validation (LOOCV) represents an extreme case in which each iteration reserves data from a single subject as the test set, while the remaining subjects are used for model development.

The evaluation began with a pilot phase conducted on the first 20 subjects drawn from the full 103-subject pool, serving as an early stopping criterion for hyperparameter screening including PSD estimation methods (average FFT vs. Welch's) and architecture selection (where convolutional layers were added incrementally). For each candidate configuration, LOOCV was early-stopped after this subset, and the combination yielding the highest average accuracy was selected and carried forward for full-scale evaluation on all 103 subjects, requiring 103 independent training cycles. Across both phases, the compact 32 × 32 input resolution was central to keeping the overall process computationally tractable. Batch Normalization (BN) layers were tested but did not consistently improve classification accuracy, so they were omitted in the final design to maintain simplicity and avoid over-regularization. This process resulted in an efficient architecture that balances model complexity and generalization, minimizing the risk of overfitting associated with excessively deep networks.

The input consists of 32 × 32 RGB-like images with either 32 × 32 × 3 channels (δ, μ, β) or 32 × 32 × 2 images (μ, β, with the blue channel zeroed), totaling 185,400 images (103 subjects × 1800 images). The CNN begins with three convolutional layers (3×3, 32 filters each) with ReLU activations, followed by max pooling. This is followed by one convolutional layer (3×3,

64 filters, ReLU), max pooling, and two convolutional layers (3×3, 128 filters, ReLU), ending with a final max pooling layer. The resulting feature map is reduced from 2048 to a 512-dimensional vector, which serves as the input to a single-layer LSTM with 128 cells. The LSTM output is connected to a fully connected layer (128 → 5), followed by a softmax activation that outputs the five classes (left fist, right fist, both fists, both feet, rest). Classification aggregates predictions across 10 sliding windows per epoch, using the final (10th) segment decision. Training was performed with a batch size of 16 over 20 epochs, using cross-entropy loss and the Adam optimizer. The overall classification workflow for the Chessboard projection approach is illustrated in Figure 14. For comparison, Table 3 summarizes the network architectures, input sizes, and key training parameters for both the Chessboard and Azimuthal projection approaches.

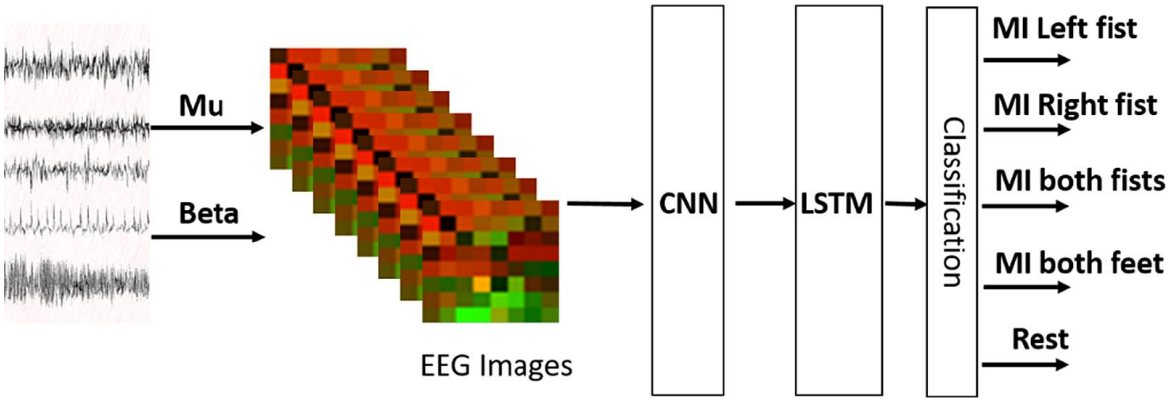

Figure 14: Classification pipeline using the Chessboard projection images representing the Mu and Beta bands. The images are processed by a CNN–LSTM architecture, which outputs five classes corresponding to motor imagery tasks (left fist, right fist, both fists, both feet) and rest.

4.3.2.2 Interpolated Azimuthal Projection

To project the 3-D electrode coordinates onto a 2-D plane, we employed the Azimuthal Equidistant Projection (AEP), with the plane tangent to the realistic head model at the vertex electrode Cz. This choice preserves the relative spatial distances of the electrodes with respect to Cz when mapping from 3-D to 2-D space.

Band-power was estimated using Welch's method, which outperformed average FFT for azimuthal projection in the pilot phase (described later). A possible reason is that azimuthal projection relies on spatial interpolation for smooth topographic maps. FFT's higher-variance estimates can thus introduce noise that propagates through interpolation and distorts the spatial representation, compared to Welch's smoothing. The average PSD values for each of the 64 electrodes within each band were encoded as color intensity specific to the band and then interpolated over a 32×32 mesh using the Clough–Tocher algorithm [179], producing smooth, unicolor, topology-preserving topomaps for each frequency band. Each segment was subsequently converted into an RGB image by mapping the topomaps to the color channels: red for Delta, green for Mu, and blue for Beta. In addition to this 3-channel configuration, a 2-channel configuration was implemented by encoding Mu and Beta in the green and red channels, respectively, while setting the blue channel to zero. This approach, particularly the generation of 2D interpolated topomaps from raw EEG, was inspired by Bashivan et al. [183], who used these spatial representations to classify mental states. Both configurations produced the same number of images for training. For each subject, the total number of images was

calculated as 30 tasks (epochs) × 10 segments per task × 6 motor imagery runs, resulting in 1,800 images per subject. Across all 103 subjects, this amounted to 1,800 × 103 = 185,400 images in total.

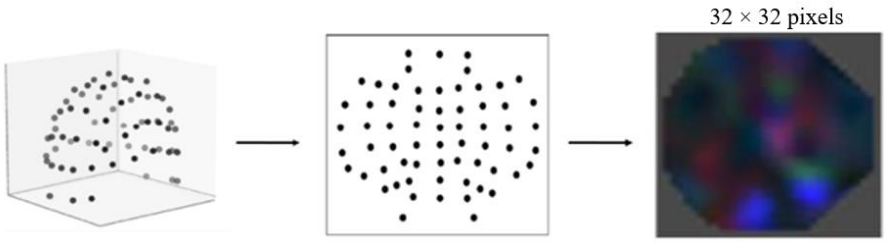

Figure 15: EEG signal-to-image transformation pipeline. Left: 3-D electrode positions on the realistic head model. Middle: 2-D electrode layout obtained using Azimuthal Equidistant Projection (AEP) with Clough–Tocher interpolation over a 32×32 mesh. Right: RGB image created by stacking the Delta (red), Mu (green), and Beta (blue) band topomaps.

Deep Learning Model

We used the same CNN–LSTM structure as for chessboard images, but with layer counts and parameters refined via pilot screening on the first 20 subjects. The input consisted of Azimuthally projected 32 × 32 × 3 RGB images (δ, μ, β) or 32 × 32 × 2 images (μ, β), totaling 185,400 images (103 subjects × 1,800 images). The CNN branch included three convolutional layers with 64 filters (3×3) each followed by ReLU activations and max pooling, two convolutional layers with 128 filters (3×3) followed by ReLU and max pooling, and two convolutional layers with 256 filters (3×3) with ReLU and a final max pooling. The CNN output was flattened from 4,096 to a 512-dimensional vector, which served as input to a single LSTM layer with 256 cells. The LSTM output was connected to a fully connected layer mapping 256 units to 5 outputs, followed by a softmax activation for classification into the five motor imagery classes. Classification aggregates predictions across 10 sliding windows per epoch, using the final (10th) segment decision. Training was performed with a batch size of 16 over 20 epochs, using cross-entropy loss and the SGD optimizer. The network architecture was designed incrementally following the same logic as the Chessboard approach, with layers added sequentially and evaluated on a subset of subjects to efficiently balance model complexity and generalization. The overall classification workflow for the Azimuthal projection approach is illustrated in Figure. 16.

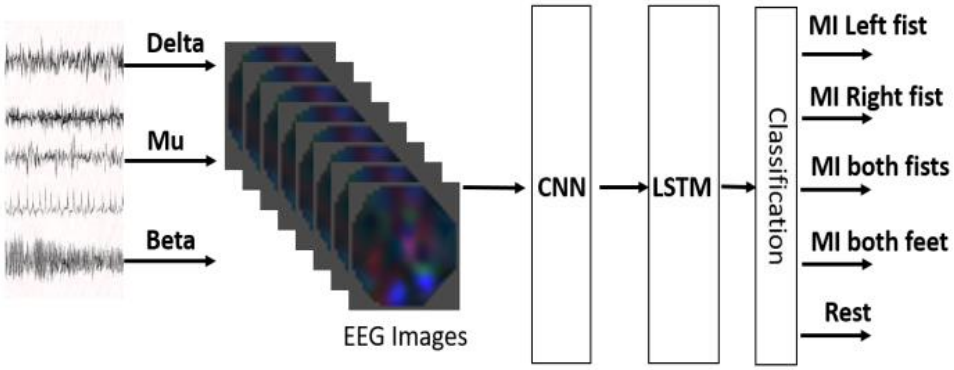

Figure 16: Classification pipeline using the Azimuthal projection images representing the Delta, Mu, and Beta bands. The images are processed by a CNN–LSTM architecture, which outputs five classes corresponding to motor imagery tasks (left fist, right fist, both fists, both feet) and rest.

For comparison, Table 3 summarizes the network architectures, input sizes, and key training parameters for both the Chessboard and Azimuthal projection approaches.

Table 3. Comparison of the Chessboard and Azimuthal projection network architectures. The table summarizes the input size, number of images, convolutional and LSTM layers, activation and pooling strategies, optimizer, loss function, batch size, and number of training epochs for both approaches.

Layer/Parameter	Chessboard Projection	Azimuthal Projection
Input Shape	32 × 32 × 3 RGB (δ, μ, β) or 32 × 32 × 2 (μ, β, Blue channel zeroed)	32 × 32 × 3 RGB (δ, μ, β) or 32 × 32 × 2 (μ, β, Blue channel zeroed)
Input Size	185,400 images (103 subjects x 1800 images)	185,400 images (103 subjects x 1800 images)
Convolutional Layers /Max Pooling	Conv (3×3, 32) → ReLU Conv (3×3, 32) → ReLU Conv (3×3, 32) → ReLU Max Pooling Conv (3×3, 64) → ReLU Max Pooling Conv (3×3, 128) → ReLU Conv (3×3, 128) → ReLU Max pooling	Conv (3×3, 64) → ReLU Conv (3×3, 64) → ReLU Conv (3×3, 64) → ReLU Max Pooling Conv (3×3, 128) → ReLU Conv (3×3, 128) → ReLU Max Pooling Conv (3×3, 256) → ReLU Conv (3×3, 256) → ReLU Max pooling
Flatten	2048 to 512	4096 to 512
LSTM	(input = 512) 1 layer, 128 cells	(input = 512) 1 layer, 256 cells
Fully Connected	128 to 5	256 to 5
Output Layer	Softmax (**5 classes**)	Softmax (**5 classes**)
Training Parameters	Batch size: 16, Epochs: 20, Loss function: Cross-Entropy, Optimizer: Adam	Batch size: 16, Epochs: 20, Loss function: Cross-Entropy, Optimizer: SGD

4.3.2.3 Results

Results for Chessboard Projection

In this study, a subject-wise LOOCV strategy was employed for performance evaluation. In each iteration, the 1,800 two-channel images (either chessboard-transformed or azimuthal-projection images) from one subject were used exclusively for testing. The images from one randomly selected subject were used for validation, and the images from the remaining 101 subjects constituted the training set. This procedure was repeated until each subject had served as the test subject once, ensuring an unbiased assessment of model generalization across subjects.The average test classification accuracy using these 2-channel images for five classes (one 'rest' and four motor imagery classes) was 68.72% for the chessboard approach and 68.13% for the azimuthal projection approach. The subject-wise classification accuracy is presented in Figure.17. Boxplot diagram is presented in Figure.18. For some subjects, such as S004 and S098, accuracy was high for both approaches. For others, one approach performed

significantly better than the other, as shown in Table 1, Table 2. These results highlight the importance of the input data representation, even when using the same number of channels.

Adding the Delta band (0.5–4 Hz) to form 3-channel images improved the performance of the azimuthal projection approach but did not improve the Chessboard projection results. This further emphasizes that the structure of the input images, not just the number of channels, strongly affects model performance.

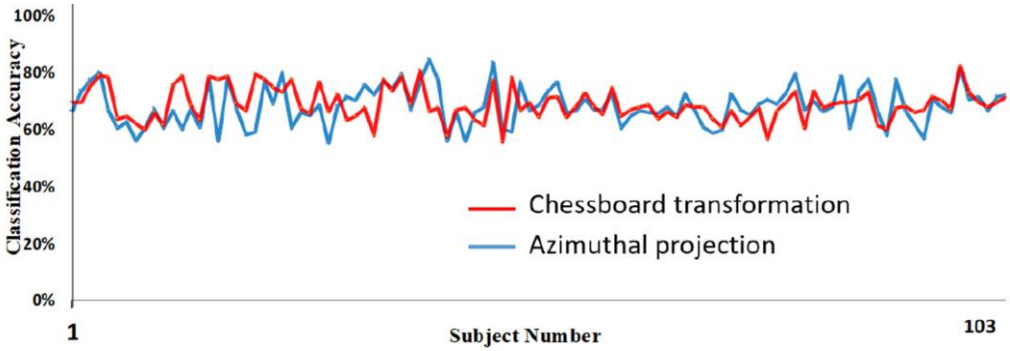

Figure 17: Classification accuracy for the (Mu and Beta) bands for each subject using leave-one-out cross-validation. The red line vertices represent the results obtained with the Chessboard projection approach, while the blue line vertices represent the results obtained with the Azimuthal projection approach.

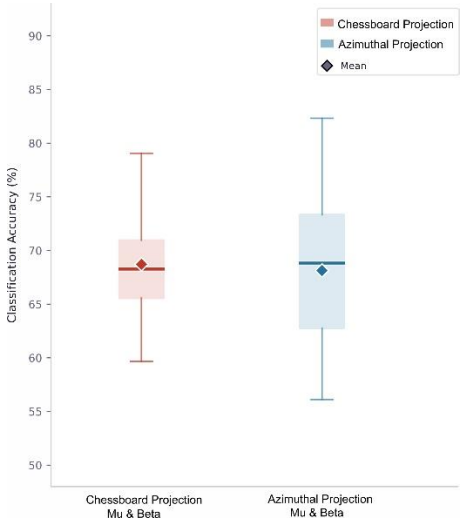

Figure 18: Boxplots of leave-one-out cross-validation classification accuracy across 103 subjects for 2-channel configuration (Mu and Beta). The red box shows Chessboard projection results, while the blue box shows Azimuthal projection results. Mean (solid line), and grand average (diamond).

The results are better compared with those of a famous baseline method for EEG classification which is Support Vector Machine (SVM) that had 64.64% classification accuracy.

Table 1. The highest performing subjects for Chessboard Projection (Mu and Beta bands).

Subject Number	Chessboard Projection accuracy	Azimuthal Projection Accuracy
S098	**82.12 %**	81.17 %
S039	**80.60 %**	77.33 %
S021	**79.34 %**	59.13 %
S004	**78.80 %**	80.04 %
S037	**78.65 %**	79.50 %

Table 2. The highest performing subjects for Azimuthal Projection (Mu and Beta bands).

Subject Number	Chessboard Projection accuracy	Azimuthal Projection Accuracy
S047	77.20 %	**83.50 %**
S098	82.12 %	**81.17 %**
S004	78.80 %	**80.04 %**
S024	73.13 %	**80.02 %**
S080	73.24 %	**79.53 %**

Results for Interpolated Azimuthal Projection

We used the Leave-One-Out-Cross-Validation (LOOCV) for evaluating the performance of the model. All the 180 trials for one subject was used for testing and the trials for another randomly chosen subject was used for validation and the rest 101 subjects trials were used as the training set.

The average test accuracy was 70.64% (using 3 bands scenario) and 68.13% (using 2 bands scenario). Figure.19 shows the classification accuracy for both scenarios with the subject number. Figure. 20 shows the boxplot diagram.

The results are promising for 5-class motor imagery EEG classification and can be compared with the state-of-the-art methods. We compared our results with a famous baseline method for EEG classification which is Support Vector Machine (SVM) and our results showed 5% better classification accuracy.

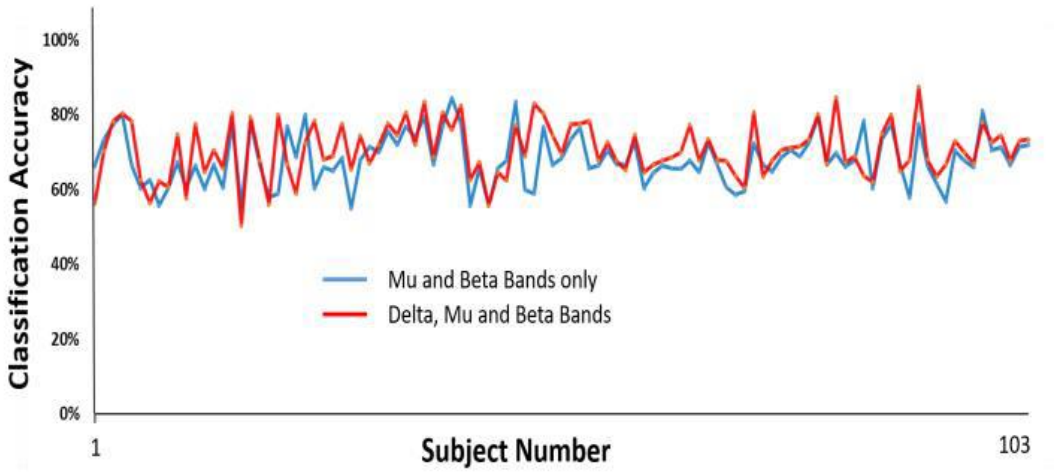

Figure 19: Classification accuracy across 103 subjects using leave-one-out cross-validation with the Azimuthal projection approach. The blue line vertices show the results for the 2-channel configuration (Mu and Beta only), while the red line vertices show the results for the 3-channel configuration (Delta, Mu, and Beta).

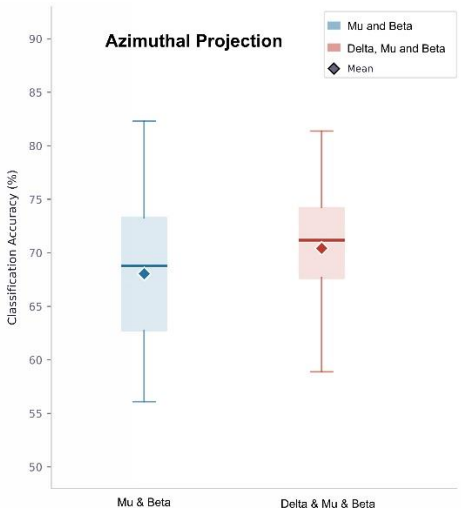

Figure 20: Boxplots of leave-one-out cross-validation classification accuracy across 103 subjects using the Azimuthal projection approach. The blue box shows results for the 2-channel configuration (Mu and Beta only), while the red box shows results for the 3-channel configuration (Delta, Mu, and Beta). Mean (solid line), grand average (diamond).

The classification performance of different models on the PhysioNet motor imagery EEG dataset is summarized in Table 4. This table compares the input features, average accuracy, and evaluation protocols for CNN–LSTM architectures using topomap images (both Chessboard and Azimuthal projections), as well as conventional classifiers such as RBF-SVM, EEGNet, and Deep ConvNet.

Table 4. classification accuracy for motor imagery EEG on PhysioNet Dataset.

Model	Input Feature	Average Accuracy %	Description
CNN + LSTM	3 Bands Topomaps (Azimuthal)	70.64	103 subjects, Leave-One-Subject-Out Cross Validation (LH, RH, BH, BF, Rest) *
	2 Bands Topomaps (Azimuthal)	68.13	
	3 Bands Topomaps (Chessboard)	68.81	-
	2 Bands Topomaps (Chessboard)	68.72	-
(RBF)SVM	3 Band PSD Features	64.64	103 subjects, 5-fold cross validation (bad epochs dropped) (LH, RH, BH, BF, Rest)
	2 Band PSD Features	60.48	
Accurate EEGNet [184]	Raw EEG, 3 s epochs	65.07	105 Subjects, Global 5-fold cross- validation (LH, RH, BH, BF)
I D ConvNet [185]	Raw EEG, 3 s epochs	58.58	109 Subjects, Global 5-fold cross-validation (LH, RH, BF, Rest)

* LH: Left Hand (Fist), RH: Right Hand, BH: Both Hands, BF: Both Feet, Rest: Rest condition

4.3.2.4 Discussion

While Mu (8–13 Hz) and Beta (13–30 Hz) rhythms dominate most motor imagery (MI) studies due to strong ERD/ERS effects, our results revealed a consistent improvement when Delta (0.5–4 Hz) features were included — but only under the Azimuthal projection. This finding motivated a closer examination of slow cortical potentials and their interaction with spatial representation in cue-based MI decoding.

1. *Neurophysiological Basis of Delta Activity*: Delta oscillations are classically associated with deep non-REM sleep and slow-wave brain states. Beyond sleep, in awake, task-engaged participants, Delta rhythms encode attention, expectation, and motor preparatory states, supporting cortical coordination and the buildup of task-relevant activity. This provides a physiological rationale for the relevance of Delta in MI decoding.

2. *Capturing Slow Cortical Dynamics*: Although a high-pass filter at 0.5 Hz suppresses artifacts, it does not eliminate all slow cortical processes. Very slow drifts may still influence the lowest frequency bins. When spectral features are extracted in sliding windows (0.4 s in our study), extremely slow changes (<0.5 Hz) are not resolved within a single window but can appear as gradual changes in Delta-band power across consecutive windows. Feeding these sequences into a unidirectional LSTM enables the network to capture slow trends, effectively modeling readiness- and anticipation-like dynamics.

3. *Readiness Potentials (RP) and CNV in Predictable Cue-Based MI*: Delta-band activity reflects slow cortical potentials such as the Readiness Potential (RP) and Contingent Negative Variation (CNV). In self-paced movement, the RP begins ~1–2 s before action (early RP) and steepens ~0.5 s before onset (late RP). In cue-based MI:

- The early RP is largely diminished, but in predictable repetitive paradigms (4 s MI / 4 s rest), participants anticipate the upcoming MI epoch. A small anticipatory RP may appear during the last ~1 s of the rest period.
- The late RP emerges during the imagery epoch, peaking ~0.5–1 s after cue onset.
- The CNV may begin near the end of the rest period and continues across the 4 s imagery epoch, reflecting sustained attention and preparation.

Even with high-pass filtering over 0.5 Hz, these effects are indirectly reflected in the temporal evolution of Delta-band power across consecutive windows. Since classification is based on 10 sliding windows per epoch with the final decision at the last segment, the LSTM can leverage cumulative Delta activity spanning anticipatory and imagery periods.

4. *Interaction with Mu and Beta Rhythms*: Slow cortical potentials are not independent of faster rhythms. *Cross-frequency coupling (CFC)* mechanisms, such as phase–amplitude modulation, allow Delta-phase dynamics to shape the amplitude and timing of Mu and Beta ERD/ERS. By including Delta, the CNN–LSTM captures both the direct contribution of slow drifts and their modulatory influence on sensorimotor rhythms, enhancing class discriminability.

5. *RGB Fusion of Frequency Bands*: Delta, Mu, and Beta are encoded as three channels of RGB topomaps. CNN filters can learn within-band spatial patterns as well as cross-band correlations, integrating slow cortical drifts with localized ERD/ERS patterns. This enables the network to capture cross-frequency interactions rather than treating each rhythm independently.

6. *Effect of Spatial Representation: Azimuthal Projection vs. Chessboard*: The benefit of Delta depends on how electrode activity is spatially represented:

- Azimuthal Projection with Clough–Tocher interpolation: Electrodes are projected onto a 2-D plane, generating smooth 32×32 maps. This preserves global spatial gradients and subtle low-frequency variations, making it ideal for identifying Delta-related RP/CNV drifts while retaining localized Mu/Beta ERD/ERS.
- Chessboard Layout: Electrodes are mapped discretely to a 4×4 grid, yielding block-like 32×32 images. While CNN filters can extract localized Mu/Beta features, smooth, widespread Delta patterns are poorly represented, limiting the capture of slow modulatory effects.

7. *Spatio-Temporal Modeling with CNN–LSTM*: Our compact VGGNet-inspired CNN extracts spatial features from topomaps. The LSTM models temporal dependencies across 10 windows per 4 s epoch, capturing:

1. Spatial topographies: readiness-like Delta distributions and localized Mu/Beta ERD/ERS.
2. Temporal dynamics: gradual buildup during the end of rest and imagery, reflecting late RP and CNV.
3. Epoch-level integration: final classification at the last segment leverages cumulative slow and fast dynamics.

Together, this hybrid CNN–LSTM framework explains why including Delta features significantly improved classification under the Azimuthal projection while offering minimal gains with the Chessboard layout. However, these conclusions require further validation.

Chapter 5

Summary

5.1 Scientific Results

Thesis Group I: Signal to Image Transformation

We propose and evaluate two end-to-end EEG signal-to-image transformation methods — Chessboard Projection and Interpolated Azimuthal Projection — that convert spectral bandpower dynamics into spatially organized 2D representations optimized for deep learning. These transformations encode neurophysiologically relevant frequency bands (Mu, Beta, and optionally Delta) into color channels, allowing the first branch of the model (CNN) to capture spatial patterns of cortical activity during motor imagery tasks. Together, they establish a unified image-based paradigm for EEG classification.

Thesis Ia: The Chessboard Projection — A Discrete, Grid-Optimized EEG-to-Image Transformation

We introduce the Chessboard Projection: a novel signal-to-image transformation that maps 64 EEG electrodes onto a rigid 32×32 pixel grid, where each electrode occupies a fixed 4×4 pixel block. Band power for Delta (0.5–4 Hz), Mu (8–13 Hz), and Beta (13–30 Hz) — estimated via Welch's PSD after 5th-order Butterworth filtering — is encoded into image channels, yielding either 3-channel maps (red for Delta, green for Mu, blue for Beta) or 2-channel maps (red for Mu, green for Beta, with the blue channel set to zero). This approach prioritizes structural regularity, computational simplicity, and alignment with convolutional receptive fields.

Thesis Ib: The Interpolated Azimuthal Projection — A Continuous, Neuroanatomically informed EEG-to-Image Transformation

We deployed the Interpolated Azimuthal Equidistant Projection: a signal-to-image transformation that projects 3D electrode positions onto a 2D plane tangent at Cz using distance-preserving geodesic mapping, followed by Clough–Tocher interpolation over a 32×32 mesh. Band power for Delta (0.5–4 Hz), Mu (8–13 Hz), and Beta (13–30 Hz) — computed via FFT after 5th-order Butterworth filtering — is encoded as 3-channel maps (red for Delta, green for Mu, blue for Beta) or 2-channel maps (red for Mu, green for Beta, blue channel set to zero). This produces smooth, continuous topomap images that preserve cortical topology and spatial gradients for CNN feature extraction.

Thesis Group II: End-to-End Spatiotemporal Learning with Hybrid CNN-LSTM

We designed a hybrid CNN–LSTM architecture tailored to the spatial properties of Chessboard and Azimuthal projections, enabling joint learning of cortical topographies and their temporal dynamics. Notably, we show that the Delta band, traditionally dismissed in motor imagery, improves accuracy only when spatially preserved through Azimuthal encoding, underscoring spatial representation as a decisive factor in EEG decoding.

***Thesis IIa*: The Architecture Co-Designed with Input Representation**

Rather than applying a fixed deep learning template, we designed the CNN-LSTM architecture in response to the structure of the input images. The CNN branch was selected to capture spatial patterns of band power distribution across the 32×32 topographic grid, whether discrete (Chessboard) or smooth (Azimuthal), leveraging its inherent strength in learning local, translation-invariant features from image-like data. The LSTM branch was introduced to model the temporal evolution of these spatial patterns across the 10 segments (0.4s each) within a motor imagery epoch — crucial for capturing the dynamic nature of cortical activation. For the discrete, grid-aligned Chessboard projections, we employed a shallower, narrower CNN to avoid overfitting and match the input's structural simplicity. For the continuous, gradient-rich Azimuthal topomaps, we scaled depth and capacity to exploit their anatomical fidelity and spatial smoothness

***Thesis IIb*: The Delta Band Effect**

Despite identical segmentation, filtering, temporal modeling, and architectural procedure, adding the Delta band (0.5–4 Hz) yielded a significant +2.5% accuracy gain — but only under Azimuthal encoding (70.64% vs. 68.13%), with no measurable improvement under Chessboard. This reveals that low-frequency cortical dynamics, long dismissed as non-informative in motor imagery, become decodable when spatially preserved through continuous, anatomy-aware interpolation. Importantly, such Delta activity overlaps with established slow cortical potentials, including the Readiness Potential (RP) and the Contingent Negative Variation (CNV), both of which reflect preparatory and anticipatory processes relevant to motor control. This result challenges the widespread assumption that low-frequency oscillations are irrelevant to motor imagery classification and suggests that even 'non-canonical' spectral components may carry latent, task-discriminative information. Future systems should treat spatial encoding not as a neutral preprocessing step, but as a decisive interface that gates what the brain reveals to the machine.

Use of AI Assistance

During the preparation of this dissertation, I used multiple generative AI tools, including the most recent versions available as of August 31, 2025, such as Grok, ChatGPT, Perplexity AI, DeepSeek, and Qwen, to assist with text generation and improve readability. All aspects of research design, analysis, and critical arguments are my own. I have reviewed and edited the outputs where appropriate and take full responsibility for the content.

Publications

[1] W. Fadel, Cs. Kollod, M. Wahdow, Y. Ibrahim, and I. Ulbert, "*Multi-Class Classification of Motor Imagery EEG Signals Using Image-Based Deep Recurrent Convolutional Neural Network*", in 2020 8th International Winter Conference on Brain-Computer Interface (BCI), 2020, pp. 1–4. doi:10.1109/BCI48061.2020.9061622.

[2] W. Fadel, M. Wahdow, Cs. Kollod, G. Marton, and I. Ulbert, "*Chessboard EEG Images Classification for BCI Systems Using Deep Neural Network*", in Bioinspired Information and Communication Technologies, Y. Chen, T. Nakano, L. Lin, M. U. Mahfuz, and W. Guo, Eds., Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering ser., vol. 329, Cham: Springer International Publishing, 2020, pp. 97–104. doi: 10.1007/978-3-030-57115-3_8.

[3] A. Adolf, Cs. Köllőd, G. Márton, W. Fadel, and I. Ulbert, *"The Effect of Processing Techniques on the Classification Accuracy of Brain-Computer Interface Systems,"* Brain Sciences, vol. 14, no. 12, p. 1272, 2024.

[4] Cs. Köllőd, A. Adolf, G. Márton, M. Wahdow, W. Fadel, and I. Ulbert, "*Closed loop BCI system for Cybathlon 2020*", Brain-Computer Interfaces, 10th vol., 2nd no., pp. 114–128, 2023. doi: 10.1080/2326263X.2023.2254463.

[5] Cs. Köllőd, A. Adolf, G. Márton, M. Wahdow, W. Fadel, and I. Ulbert, TTK dataset - 4 class Motor-Imagery EEG, 2022, https://repo.researchdata.hu/dataset.xhtml?persistentId=hdl:21.15109/CONCORDA/UOQQVK.

[6] M.Wahdow, M. Alnaanah, W. Fadel, A. Adolf, Cs. Kollod, and I. Ulbert, "*Multi frequency band fusion method for EEG signal classification*", Signal, Image and Video Processing, 2022. doi: 10.1007/s11760-022-02399-6.

[7] W. Fadel, R. Tóth, R. Fiáth, I. Ulbert, and Z. Somogyvári, "*Optimal inter-electrode distances for spike sorting in different brain regions*", (Poster, HunDoc 2024, Hungary)

[8] W. Fadel, I. Ulbert, "*Convolutional Neural Network Classifier for BCI systems using Motor Imagery EEG Signals-to-Image Transformation*" (Poster for CYBATHLON Symposium 2020, SWISS Arena, Switzerland).

Bibliography

[1] E. Niedermeyer and F. L. da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2004.

[2] M. A. B. Brazier, *A History of the Electrical Activity of the Brain: The First Half-Century*. Washington, DC: Spartan Books, 1961.

[3] L. F. Haas, "Hans Berger (1873–1941), Richard Caton (1842–1926), and electroencephalography," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 74, no. 1, p. 9, 2003.

[4] P. Gloor, "The role of the cerebral cortex in the genesis of the EEG," in *Physiological Basis of Epilepsy*, H. H. Jasper, A. A. Ward, and A. Pope, Eds. Springfield, IL: Charles C Thomas, 1969, pp. 1–20.

[5] V. Pravdich-Neminsky, "Ein Versuch der registrierung der elektrischen Gehirnerscheinungen," Archiv für die gesamte Physiologie des Menschen und der Tiere, vol. 159, pp. 1–12, 1913.

[6] H. Berger, "Über das Elektrenkephalogramm des Menschen," Archiv für Psychiatrie und Nervenkrankheiten, vol. 87, no. 1, pp. 527–570, 1929.

[7] E. D. Adrian and B. H. C. Matthews, "The interpretation of electrical potentials from the surface of the human brain," The Journal of Physiology, vol. 82, no. 1, pp. 7–10, 1934.

[8] B. J. Fisch, "Hans Berger: From psychic energy to the EEG," Epilepsy & Behavior, vol. 1, no. 3, pp. 216–224, 1999.

[9] F. A. Gibbs, H. Davis, and W. G. Lennox, "The electroencephalogram in epilepsy and in conditions of impaired consciousness," Archives of Neurology and Psychiatry, vol. 34, no. 6, pp. 1133–1155, 1935.

[10] A. Aserinsky and N. Kleitman, "Regularly occurring periods of eye motility, and concomitant phenomena, during sleep," Science, vol. 118, no. 3062, pp. 273–274, 1953.

[11] R. Cooper, J. W. Osselton, and J. C. Shaw, EEG Technology, 2nd ed. London: Butterworths, 1969.

[12] H. H. Jasper, "The ten-twenty electrode system of the International Federation," Electroencephalography and Clinical Neurophysiology, vol. 10, pp. 371–375, 1958.

[13] E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth, Principles of Neural Science, 5th ed. New York, NY: McGraw-Hill, 2013.

[14] P. L. Nunez and R. Srinivasan, Electric Fields of the Brain: The Neurophysics of EEG, 2nd ed. Oxford: Oxford University Press, 2006.

[15] D. Berger and M. Januszewski, "6 incredible images of the human brain built with the help of Google's AI," Google Research Blog, May 9, 2024. [Online]. Available: https://blog.google/technology/research/google-ai-research-new-images-human-brain

[16] G. Buzsáki, N. Logothetis, and W. Singer, "Scaling brain size, keeping timing: Evolutionary preservation of brain rhythms," Neuron, vol. 80, no. 3, pp. 751–764, 2012. doi: 10.1016/j.neuron.2013.10.002

[17] G. Buzsáki, Rhythms of the Brain. Oxford: Oxford University Press, 2006.

[18] J. Szentágothai, "The neuron network of the cerebral cortex: A functional interpretation," Proceedings of the Royal Society of London. Series B, Biological Sciences, vol. 221, no. 1222, pp. 255–272, 1983.

[19] Epomedicine, "Brodmann Areas and Lesions," Epomedicine, Jul. 31, 2016. [Online]. Available: https://epomedicine.com/medical-students/brodmann-areas-lesions/

[20] G. Buzsáki, "Theta oscillations in the hippocampus," Neuron, vol. 33, no. 3, pp. 325–340, 2002.

[21] G. Buzsáki and X. J. Wang, "Mechanisms of gamma oscillations," Annual Review of Neuroscience, vol. 35, pp. 203–225, 2012.

[22] S. W. Hughes and V. Crunelli, "Thalamic mechanisms of EEG alpha rhythms and their pathological implications," The Neuroscientist, vol. 11, no. 4, pp. 357–372, 2005.

[23] M. Steriade, D. A. McCormick, and T. J. Sejnowski, "Thalamocortical oscillations in the sleeping and aroused brain," Science, vol. 262, no. 5134, pp. 679–685, 1993.

[24] M. Steriade, "Grouping of brain rhythms in corticothalamic systems," Neuroscience, vol. 137, no. 4, pp. 1087–1106, 2006.

[25] E. Neher and B. Sakmann, "Single-channel currents recorded from membrane of denervated frog muscle fibres," Nature, vol. 260, no. 5554, pp. 799–802, 1976.

[26] G. Buzsáki, "Large-scale recording of neuronal ensembles," Nature Neuroscience, vol. 7, no. 5, pp. 446–451, 2004.

[27] J.-P. Lachaux, D. Rudrauf, and P. Kahane, "Intracranial EEG and human brain mapping," Journal of Physiology-Paris, vol. 97, no. 4–6, pp. 613–628, 2003.

[28] S. Baillet, "Magnetoencephalography for brain electrophysiology and imaging," Nature Neuroscience, vol. 20, no. 3, pp. 327–339, 2017.

[29] L. R. Hochberg et al., “Neuronal ensemble control of prosthetic devices by a human with tetraplegia,” Nature, vol. 442, no. 7099, pp. 164–171, 2006.

[30] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, “Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of the working human brain,” Reviews of Modern Physics, vol. 65, no. 2, pp. 413–497, 1993.

[31] 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, 2012.

[32] H. Obrig and A. Villringer, “Beyond the visible — imaging the human brain with light,” Journal of Cerebral Blood Flow & Metabolism, vol. 23, no. 1, pp. 1–18, 2003.

[33] N. K. Logothetis, “What we can do and what we cannot do with fMRI,” Nature, vol. 453, no. 7197, pp. 869–878, 2008.

[34] S. A. Huettel, A. W. Song, and G. McCarthy, Functional Magnetic Resonance Imaging, 2nd ed. Sunderland, MA: Sinauer Associates, 2004.

[35] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, “Brain–computer interfaces for communication and control,” Clinical Neurophysiology, vol. 113, no. 6, pp. 767–791, Jun. 2002.

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

[37] F. Lotte et al., “A review of classification algorithms for EEG-based brain–computer interfaces,” Journal of Neural Engineering, vol. 15, no. 3, p. 031001, Jun. 2018.

[38] E. 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, 1988.

[39] J. Polich, “Updating P300: An integrative theory of P3a and P3b,” Clinical Neurophysiology, vol. 118, no. 10, pp. 2128–2148, 2007.

[40] J. Middendorf, G. McMillan, G. Calhoun, and L. Jones, “Brain-computer interfaces based on the steady-state visual-evoked response,” IEEE Transactions on Rehabilitation Engineering, vol. 8, no. 2, pp. 211–214, Jun. 2000.

[41] Y. Zhang, B. Xu, and C. Guan, “Frequency recognition in SSVEP-based BCI using canonical correlation analysis,” Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4133–4136, 2010.

[42] N. Birbaumer, M. Elbert, T. Canavan, and E. Rockstroh, “Slow potentials of the cerebral cortex and behavior,” Physiological Reviews, vol. 70, no. 1, pp. 1–41, 1990.

[43] A. Kübler, N. Birbaumer, T. Hinterberger, B. Weiskopf, and B. Neumann, "Brain–computer communication: Unlocking the locked in," Psychological Bulletin, vol. 127, no. 3, pp. 358–375, 2001.

[44] P. W. Ferrez and J. R. Millán, "Error-related EEG potentials generated during simulated brain–computer interaction," IEEE Transactions on Biomedical Engineering, vol. 55, no. 3, pp. 923–929, Mar. 2008.

[45] R. Chavarriaga and J. 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, Aug. 2010.

[46] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," Proceedings of the IEEE, vol. 89, no. 7, pp. 1123–1134, Jul. 2001.

[47] M. Jeannerod, Mental imagery in the motor context, Neuropsychologia, vol. 33, pp. 1419–1432, 1994.

[48] R. Lotze and U. Halsband, "Motor imagery," J. Physiol. Paris, vol. 99, no. 4–6, pp. 386–395, 2006.

[49] M. Kasess, R. Windischberger, U. Cunnington, E. Lanzenberger, L. Pezawas, and R. Moser, "The suppressive influence of SMA on M1 during motor imagery: Evidence from fMRI," *NeuroImage*, vol. 40, no. 1, pp. 174–182, 2008.

[50] R. Lotze et al., "Activation of cortical and cerebellar motor areas during executed and imagined hand movements," Cogn. Brain Res., vol. 7, pp. 235–243, 1999.

[51] F. Lopes da Silva, EEG: Basic principles, clinical applications, and related fields, 2nd ed., 2006.

[52] C. Neuper and G. Pfurtscheller, "Motor imagery and ERD/ERS," in Progress in Brain Research, vol. 159, 2001, pp. 203–211.

[53] G. Pfurtscheller and F. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," Clin. Neurophysiol., vol. 110, no. 11, pp. 1842–1857, 1999.

[54] A. Guillot, C. Collet, N. Nguyen, M. Malouin, C. Richards, and J. Doyon, "Imagery in motor learning," *NeuroImage*, vol. 47, pp. 125–132, 2009.

[55] B. Hétu et al., "The neural network of motor imagery: An ALE meta-analysis," Neurosci. Biobehav. Rev., vol. 37, pp. 930–949, 2013.

[56] S. Grèzes and J. Decety, "Functional anatomy of execution, mental simulation, observation, and verb generation of actions," Hum. Brain Mapp., vol. 12, pp. 1–19, 2001.

[57] A. Solodkin, P. Hlustik, E. E. Chen, and S. L. Small, "Fine modulation in network activation during motor imagery," *Cereb. Cortex*, vol. 14, pp. 1246–1255, 2004.

[58] C. Neuper, R. Scherer, S. Wriessnegger, and G. Pfurtscheller, “EEG-based neurofeedback and motor imagery: Methods and clinical applications,” *Clin. Neurophysiol.*, vol. 116, pp. 2380–2390, 2005.

[59] C. Neuper, R. Wörtz, and G. Pfurtscheller, “ERD/ERS patterns reflecting sensorimotor activation and deactivation,” Brain Topography, vol. 8, no. 4, pp. 203–210, 2006.

[60] C. Neuper, G. Pfurtscheller, and R. Wörtz, “Imagery of motor actions: Differential effects of kinesthetic and visual–motor mode of imagery in single-trial EEG,” Cognitive Brain Research, vol. 25, no. 3, pp. 668–677, 2005.

[61] H. Jasper and W. Penfield, “Electrocorticograms in man: Effect of voluntary movement upon the electrical activity of the precentral gyrus,” Arch. Psychiatr. Nervenkr., vol. 183, pp. 163–174, 1949.

[62] J. Arroyo, L. Lesser, M. Gordon, and M. Uematsu, “Functional significance of the mu rhythm of human cortex: An electrocorticographic study with subdural electrodes,” Electroencephalogr. Clin. Neurophysiol., vol. 87, pp. 76–87, 1993.

[63] N. E. Crone, D. L. Miglioretti, B. Gordon, and R. P. Lesser, “Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. I. Alpha and beta event-related desynchronization,” *Brain*, vol. 121, pp. 2271–2299, 1998.

[64] R. Salmelin and R. Hari, “Spatiotemporal characteristics of sensorimotor neuromagnetic rhythms related to thumb movement,” Neuroscience, vol. 60, pp. 537–550, 1994.

[65] R. Salmelin, M. Schnitzler, R. Schmitz, and R. Freund, “Functional segregation of movement-related rhythmic activity in the human brain,” NeuroImage, vol. 2, pp. 237–243, 1995.

[66] G. Pfurtscheller, C. Neuper, A. Andrew, and W. Edlinger, “Foot and hand area mu rhythms,” Int. J. Psychophysiol., vol. 26, pp. 121–135, 1997.

[67] G. Pfurtscheller, A. Stancák, and C. Neuper, “Post-movement beta synchronization in EEG,” *Electroencephalography and Clinical Neurophysiology*, vol. 96, pp. 197–207, 1995.

[68] R. Salenius, M. Kajola, R. Thompson, K. Kosslyn, and R. Hari, “Cortical control of human motor function: The mu rhythm and its modulation by movement,” *Experimental Brain Research*, vol. 114, pp. 557–566, 1997.

[69] M. Schnitzler, S. Salenius, R. Salmelin, V. Jousmäki, and R. Hari, “Synchronization in the human sensorimotor system,” *Science*, vol. 274, pp. 892–895, 1996.

[70] M. Alegre, I. G. Gurtubay, J. Labarga, M. Iriarte, C. Malanda, and J. Artieda, “Movement-related changes in oscillatory activity in the human motor cortex: Reafference and beta modulation,” *NeuroReport*, vol. 13, pp. 881–884, 2002.

[71] D. Tan, T. W. W. Chou, D. M. N. Guo, and P. Brown, “Post-movement beta rebound and predictive coding in motor cortex,” *J. Neurosci.*, vol. 36, pp. 839–848, 2016.

[72] S. Shin, N. Law, R. Tsutsumi, and T. Yamawaki, “Beta oscillations reflect network resetting in sensorimotor cortex,” *Cereb. Cortex*, vol. 27, pp. 2982–2993, 2017.

[73] R. Chen *et al.*, “TMS-induced modulation of motor cortex excitability,” *Neurology*, vol. 50, pp. 1340–1345, 1998.

[74] R. Chen, A. Lozano, M. Ashby, and M. Hallett, “Suppression of somatosensory-evoked potentials by TMS over motor cortex,” *Exp. Brain Res.*, vol. 129, pp. 167–175, 1999.

[75] P. Jenkinson and P. Brown, “New insights into the relationship between dopamine, Beta oscillations and motor function,” Trends Neurosci., vol. 34, pp. 611–618, 2011.

[76] T. Ikeda, M. Lüders, H. Burgess, and H. Shibasaki, “Foot area representation in human sensorimotor cortex: ECoG evidence,” *Neurosci. Lett.*, vol. 139, pp. 157–162, 1992.

[77] T. Ohara, E. Ikeda, M. Kunieda, and T. Shibasaki, “Functional mapping of foot and leg sensorimotor cortex,” *J. Neurosci.*, vol. 20, pp. 4851–4858, 2000.

[78] G. Pfurtscheller, C. Neuper, A. Stancák, and M. Lopes da Silva, “Beta rebound and motor imagery: ECoG studies,” *Clin. Neurophysiol.*, vol. 111, pp. 1463–1471, 2000.

[79] T. Luu, A. Nakayama, M. Kambara, and Y. Koike, “PMBR as a movement termination signal in BCIs,” *J. Neural Eng.*, vol. 13, p. 046018, 2016.

[80] H. Park, S. Kim, J. Kim, and M. Hwang, “Post-movement beta rebound as a prognostic marker in stroke rehabilitation,” *Neurorehabil. Neural Repair*, vol. 33, pp. 567–578, 2019.

[81] H. Deecke and H. H. Kornhuber, “An electrical sign of participation of the mesial ‘supplementary’ motor cortex in human voluntary finger movement,” Brain Res., vol. 159, pp. 473–476, 1978.

[82] T. Shibasaki and M. Hallett, “What is the Bereitschaftspotential?” Clin. Neurophysiol., vol. 117, no. 11, pp. 2341–2356, 2006.

[83] J. Jankelowitz and J. G. Colebatch, “Movement-related potentials associated with self-paced, cued and imagined arm movements,” Exp. Brain Res., vol. 147, no. 1, pp. 98–107, 2002.

[84] F. Gerardin et al., “Partially overlapping neural networks for real and imagined hand movements,” Cereb. Cortex, vol. 10, pp. 1093–1104, 2000.

[85] C. Cunnington, P. Iansek, and R. Bradshaw, “Movement-related potentials in motor imagery: Lateralized readiness potentials,” NeuroImage, vol. 28, pp. 296–308, 2005.

[86] S. Munzert, M. Lorey, and J. Zentgraf, “Cognitive motor processes: The role of motor imagery in the study of motor representations,” Brain Res. Rev., vol. 60, pp. 306–326, 2009.

[87] M. Bareš, I. Lungu, T. Liu, and M. Hallett, “The role of the supplementary motor area in motor imagery: A functional MRI study,” Brain Research, vol. 1131, no. 1, pp. 174–178, 2007.

[88] H. Walter, H. Cooper, R. Aldridge, W. McCallum, and L. Winter, “Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain,” Nature, vol. 203, pp. 380–384, 1964.

[89] R. Tecce, “Contingent negative variation (CNV) and psychological processes in humans,” Psychol. Bull., vol. 77, pp. 73–108, 1972.

[90] A. L. Goldberger et al., “PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals,” Circulation, vol. 101, pp. e215–e220, 2000.

[91] J. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, “BCI2000: A general-purpose brain-computer interface (BCI) system,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, 2004.

[92] M. Kropp, J. Gerloff, A. Knab, H. Münchau, and J. H. Hallett, “Motor imagery of sequential finger movements in humans: A high-resolution EEG study,” NeuroImage, vol. 12, pp. 310–320, 2000.

[93] Y. Bai, G. Lin, B. Wang, and F. Zhang, “ERP-based brain-computer interfaces for motor imagery: A review,” Front. Comput. Neurosci., vol. 5, p. 80, 2011.

[94] K. Williams, P. Cumming, J. Ntoumanis, R. Nordin-Bates, and C. Ramsey, “The Movement Imagery Questionnaire–3 (MIQ-3): Its psychometric properties,” J. Sports Sci., vol. 30, no. 10, pp. 1110–1119, 2012.

[95] D. F. Marks, “Visual imagery differences in the recall of pictures,” Br. J. Psychol., vol. 64, pp. 17–24, 1973.

[96] A. Zeman, F. Milton, S. Della Sala, D. Dewar, and M. Smith, “Lives without imagery – Congenital aphantasia,” *Cortex*, vol. 105, pp. 53–60, 2020.

[97] A. Zeman, M. Dewar, and S. Della Sala, “Deficits in mental imagery: Congenital aphantasia,” Cortex, vol. 73, pp. 378–380, 2015.

[98] E. Dupont, L. Martin, and R. Brown, “Motor imagery deficits in individuals with aphantasia,” *Neurosci. Lett.*, vol. 819, p. 137360, 2024.

[99] J. Speed, A. White, and C. Jones, “Cortical excitability during motor imagery in aphantasia,” *Front. Hum. Neurosci.*, vol. 19, p. 123, 2025.

[100] S. Kosslyn, “Image and mind,” Harvard University Press, 1980.

[101] M. Cui, A. J. Jeter, E. A. Yang, K. Montague, and J. L. Eagleman, “Vividness of mental imagery: Individual variability and neural correlates,” Cereb. Cortex, vol. 17, pp. 221–231, 2007.

[102] T. Pearson, R. Clifford, and P. Tong, “The functional impact of imagery duration on neural activation,” Nat. Neurosci., vol. 18, pp. 618–622, 2015.

[103] C. Hall, J. Buckolz, and R. Fishburne, “Imagery and motor skill learning,” Res. Q. Exerc. Sport, vol. 56, no. 2, pp. 123–127, 1985.

[104] S. Guillot and C. Collet, “Duration of mentally simulated movement: A review,” Cogn. Brain Res., vol. 24, pp. 189–200, 2005.

[105] R. Hardy and N. Callow, “A cognitive approach to motor imagery: Perspective matters,” J. Sports Sci., vol. 17, no. 4, pp. 275–285, 1999.

[106] J. Decety, Do imagined and executed actions share the same neural substrate?, Cogn. Brain Res., vol. 3, no. 2, pp. 87–93, 1996.

[107] E. A. Holmes and A. Mathews, “Mental imagery in emotion and emotional disorders,” Clin. Psychol. Rev., vol. 30, no. 3, pp. 349–362, 2010.

[108] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller, “Optimizing spatial filters for robust EEG single-trial analysis,” IEEE Signal Process. Mag., vol. 25, no. 1, pp. 41–56, 2010.

[109] C. Vidaurre and B. Blankertz, “Towards a cure for BCI illiteracy,” Brain Topogr., vol. 23, pp. 194–198, 2010.

[110] B. Allison and C. Neuper, “Could anyone use a BCI?” in Brain–Computer Interfaces: Human–Computer Interaction Series, D. Tan and A. Nijholt, Eds. London: Springer, 2010, pp. 35–54.

[111] F. Lebon, A. Guillot, and C. Collet, “Increased muscle activation following motor imagery during the rehabilitation of the upper limb in hemiplegic patients,” Int. J. Psychophysiol., vol. 84, pp. 60–68, 2012.

[112] H. H. Jasper, “The ten-twenty electrode system of the International Federation,” Electroencephalogr. Clin. Neurophysiol., vol. 10, pp. 371–375, 1958, doi: 10.1016/0013-4694(58)90058-8.

[113] International Federation of Societies for Electroencephalography, “Standardization of techniques for recording the electroencephalogram. Report of the Committee on Nomenclature and Statistics,” Electroencephalogr. Clin. Neurophysiol., vol. 10, no. 3, pp. 371–375, 1958.

[114] G. E. Chatrian, E. Lettich, and P. L. Nelson, “Ten percent electrode system for topographic studies of spontaneous and evoked EEG activity,” Am. J. Electroencephalogr. Technol., vol. 25, no. 2, pp. 83–92, 1985, doi: 10.1093/ajet/25.2.83.

[115] V. Jurcak, D. Tsuzuki, and I. Dan, "10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems," NeuroImage, vol. 34, no. 4, pp. 1600–1611, 2007, doi: 10.1016/j.neuroimage.2006.09.024.

[116] G. H. Klem, H. Lüders, H. Jasper, and C. Elger, "The international 10–5 system for high-density EEG recordings," *Clin. Neurophysiol.*, vol. 110, Suppl. 1, pp. S15–S16, 1999.

[117] R. Oostenveld, Improving EEG Source Analysis Using Prior Knowledge, Doctoral thesis, Radboud University, Nijmegen, 2003. ISBN 90-9016991-1.

[118] W. C. Ferree, M. Luu, S. Russell, and G. Tucker, "Scalp electrode impedance, infection risk, and EEG data quality," Clin. Neurophysiol., vol. 112, no. 3, pp. 536–544, 2001.

[119] M. Widmann, E. Schröger, and B. Maess, "Digital filter design for electrophysiological data – a practical approach," J. Neurosci. Methods, vol. 250, pp. 34–46, 2015.

[120] M. X. Cohen, Analyzing Neural Time Series Data: Theory and Practice, 2nd ed., Cambridge, MA: MIT Press, 2014.

[121] H. Nolan, R. Whelan, and R. Reilly, "FASTER: Fully automated statistical thresholding for EEG artifact rejection," J. Neurosci. Methods, vol. 192, no. 1, pp. 152–162, 2010, doi: 10.1016/j.jneumeth.2010.07.015.

[122] S. Makeig, "Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones," Electroencephalogr. Clin. Neurophysiol., vol. 86, no. 4, pp. 283–293, 1993, doi: 10.1016/0013-4694(93)90029-6.

[123] J. Dien, "Issues in the application of the average reference: Review, critiques, and recommendations," Behavior Research Methods, Instruments, & Computers, vol. 30, no. 1, pp. 34–43, 1998.

[124] J. Kayser and C. E. Tenke, "On the benefits of using surface Laplacian (Current Source Density) methodology in EEG/ERP research," International Journal of Psychophysiology, vol. 97, no. 3, pp. 171–173, 2015.

[125] D. Yao, "A method to standardize a reference of scalp EEG recordings to a point at infinity," Physiological Measurement, vol. 22, no. 4, pp. 693–711, 2001.

[126] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," Human Brain Mapping, vol. 38, no. 11, pp. 5391–5420, 2017.

[127] X. Zhang, Y. Li, C. Wang, and Z. Zhang, "Data augmentation for EEG-based brain-computer interface: A review," IEEE Access, vol. 8, pp. 221859–221877, 2020.

[128] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T.-H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," Journal of Neural Engineering, vol. 16, no. 5, p. 051001, 2019.

[129] A. Subasi and M. I. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," Expert Systems with Applications, vol. 37, no. 12, pp. 8659–8666, 2010.

[130] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter bank common spatial pattern (FBCSP) in brain-computer interface," in Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN), Hong Kong, 2008, pp. 2390–2397.

[131] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," IEEE Transactions on Audio and Electroacoustics, vol. 15, no. 2, pp. 70–73, 1967.

[132] H. Ramoser, J. Müller-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, Dec. 2000.

[133] T. 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.

[134] Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, and A. Cichocki, "Sparse Bayesian learning for obtaining sparse CSPs in motor-imagery BCI," IEEE Transactions on Neural Networks and Learning Systems, vol. 27, no. 3, pp. 621–632, Mar. 2016.

[135] 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, Feb. 2020.

[136] R. A. Fisher, "The use of multiple measurements in taxonomic problems," Annals of Eugenics, vol. 7, no. 2, pp. 179–188, 1936.

[137] G. Dornhege, B. Blankertz, G. Curio, and K.-R. Müller, "Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms," IEEE Transactions on Biomedical Engineering, vol. 51, no. 6, pp. 993–1002, Jun. 2004.

[138] B. Blankertz, K.-R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlögl, G. Pfurtscheller, J. R. Millán, M. Schröder, and N. Birbaumer, "The BCI competition III: Validating alternative approaches to actual BCI problems," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 14, no. 2, pp. 153–159, Jun. 2006.

[139] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," Journal of Multivariate Analysis, vol. 88, no. 2, pp. 365–411, 2004.

[140] Y. Liu, D. Zhang, and Y. Wang, “Classification of EEG signals using Heteroscedastic Linear Discriminant Analysis,” in Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 2007, pp. 1375–1378.

[141] K. Fukunaga, Introduction to Statistical Pattern Recognition, 2nd ed. San Diego, CA: Academic Press, 1990.

[142] V. N. Vapnik, The Nature of Statistical Learning Theory, New York, NY: Springer, 1995.

[143] C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[144] B. Schölkopf and A. J. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond, Cambridge, MA: MIT Press, 2002.

[145] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods, Cambridge: Cambridge Univ. Press, 2000.

[146] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, “A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update,” J. Neural Eng., vol. 15, no. 3, p. 031005, 2018.

[147] M. Congedo, A. Barachant, and R. Bhatia, “Riemannian geometry for EEG-based brain–computer interfaces; a primer and a review,” Brain-Comput. Interfaces, vol. 4, no. 3, pp. 155–174, 2017.

[148] M. Congedo, A. Barachant, and K. Kharati, “Classification of covariance matrices using a Riemannian-based kernel for BCI applications,” IEEE Trans. Signal Process., vol. 65, no. 9, pp. 2211–2220, 2017.

[149] M. Wairagkar, “Motor Imagery based Brain–Computer Interface (BCI) using Artificial Neural Networks classifiers,” Ph.D. dissertation, Univ. of Reading, Reading, U.K., 2016.

[150] W. Abbas and N. Khan, “DeepMI: deep learning for multiclass motor imagery classification,” in Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), 2018, pp. 219–222.

[151] W. Qiao and J. Bi, “IncepCNN-BGRU for motor imagery EEG classification,” Expert Systems with Applications, vol. 201, 116789, 2022.

[152] Xu, J., Zheng, H., Wang, J., Li, D., & Fang, X. “Recognition of EEG Signal Motor Imagery Intention Based on Deep Multi-View Feature Learning.” *Sensors*, vol. 20, no. 3496, 2020.

[153] S. Chaudhary, S. Taran, V. Bajaj, and A. Sengur, “Convolutional neural network based approach towards motor imagery tasks EEG signals classification,” IEEE Sensors J., vol. 19, no. 12, pp. 4494–4502, 2019.

[154] Z. Wang and T. Oates, “Encoding time series as images for visual inspection and classification using tiled convolutional neural networks,” in Proc. AAAI Workshop on Learning Representations for Time Series (AAAI), 2015, pp. 40–46.

[155] Z. Zhang, M. Cui, and L. Liu, “Time series classification using Gramian angular fields and Markov transition fields,” IEEE Access, vol. 8, pp. 123456–123467, 2020.

[156] H. Li, J. Zhang, X. Qin, and Y. Wang, “Recurrence plots with graph neural networks for EEG-based motor imagery classification,” Frontiers in Neuroscience, vol. 16, art. 1097660, 2022.

[157] J. Chen and Z. Wang, “Angle-Amplitude transformation for EEG signal encoding in motor imagery tasks,” IEEE Trans. Biomed. Eng., vol. 70, no. 4, pp. 1123–1135, 2023.

[158] Z. Wan, M. Li, S. Liu, J. Huang, H. Tan, and W. Duan, “EEGFormer: A transformer-based brain activity classification method using EEG signals,” Frontiers in Neuroscience, vol. 17, art. 1148855, Mar. 2023.

[159] L. Wang, Y. Xie, and K. Zhang, “ST-ViT: Spatial-Temporal Vision Transformer for EEG topographic map classification,” IEEE Trans. Pattern Anal. Mach. Intell., 2024.

[160] H. Ramoser, J. Müller-Gerking, and P. A. Schönle, “Optimal spatial filtering of single trial EEG during imagined hand movement,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–448, Dec. 2000.

[161] H. Lu, C. L. Xu, and C. L. Guan, “Regularized common spatial pattern with aggregation for EEG classification in small-sample settings,” *Neurocomputing*, vol. 74, no. 17, pp. 4061–4067, Dec. 2011.

[162] X. Jiang, L. Meng, X. Chen, Y. Xu, and D. Wu, “CSP-Net: Common Spatial Pattern Empowered Neural Networks for EEG-Based Motor Imagery Classification,” Knowledge-Based Systems, vol. 305, article 112668, Dec. 2024.

[163] H. Wang, J. Tian, S. Li, H. Zhao, F. Wu, and X. Li, "Structure-conditioned adversarial learning for unsupervised domain adaptation," Neurocomputing, vol. 497, pp. 216–226, Aug. 2022.

[164] K. T. Kim, S. J. Lee, D. J. Kim, and H. Kim, "Subject-transfer with subject-specific fine-tuning based on multi-model CNN for motor imagery brain-computer interface," in Proc. 11th Int. Winter Conf. Brain-Computer Interface (BCI), Gangwon, South Korea, pp. 1–5, Feb. 2023.

[165] L. Fan, B. Chen, X. Zeng, J. Zhou, and X. Zhang, "Knowledge-enhanced meta-transfer learning for few-shot ECG signal classification," Expert Syst. Appl., vol. 252, Art. no. 124764, Oct. 2024.

[166] M. Jiao, S. Yang, X. Xian, N. Fotedar, and F. Liu, “Multi-Modal Electrophysiological Source Imaging With Attention Neural Networks Based on Deep

Fusion of EEG and MEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 32, pp. 2492–2502, 2024.

[167] Y. Wu, P. Cao, M. Xu, Y. Zhang, X. Lian, and C. Yu, "Adaptive GCN and Bi-GRU-Based Dual Branch for Motor Imagery EEG Decoding," Sensors, vol. 25, no. 4, p. 1147, Feb. 2025.

[168] S. Torma and L. Szegletes, "Generative modeling and augmentation of EEG signals using improved diffusion probabilistic models," J. Neural Eng., vol. 22, no. 1, Jan. 2025

[169] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[170] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiol.*, vol. 160, no. 1, pp. 106–154, Jan. 1962.

[171] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biol. Cybern.*, vol. 36, no. 4, pp. 193–202, 1980.

[172] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.

[173] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1097–1105.

[174] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," presented at the Int. Conf. Learn. Represent. (ICLR), San Diego, CA, USA, May 7–9, 2015.

[175] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 27–30, 2016, pp. 770–778.

[176] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.

[177] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[178] F. Perrin, J. Pernier, O. Bertrand, and J. F. Echallier, "Spherical splines for scalp potential and current density mapping," *Electroencephalogr. Clin. Neurophysiol.*, vol. 72, no. 2, pp. 184–187, Feb. 1989.

[179] R. Franke, "Smooth interpolation of scattered data by local polynomials," *ACM Trans. Math. Softw.*, vol. 8, no. 2, pp. 119–143, Jun. 1982.

[180] D. F. Watson, *Contouring: A Guide to the Analysis and Display of Spatial Data*. Oxford, U.K.: Pergamon, 1992.

[181] N. Cressie, *Statistics for Spatial Data*. New York, NY, USA: Wiley, 1993.

[182] A. Loboda, A. Margineanu, G. Rotariu, and A. M. Lazar, "Discrimination of EEG-based motor imagery tasks by means of a simple phase information method," Int. J. Adv. Res. Artif. Intell., vol. 3, no. 10, 2014.

[183] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent convolutional neural networks," in Proc. Int. Conf. Learn. Represent. (ICLR), 2016.

[184] X. Wang, Y. Yang, and Y. Zhang, "Accurate EEGNet-based motor imagery classification with minimal preprocessing," *Journal of Neural Engineering*, vol. 20, no. 4, p. 046012, 2023.

[185] H. Dose, J. S. Møller, H. K. Iversen, and S. Puthusserypady, "An end-to-end deep learning approach to MI-EEG signal classification for BCI," *Frontiers in Neuroscience*, vol. 12, p. 711, 2018.