

NeuroPath: Practically Adopting Motor Imagery Decoding through EEG Signals

Jiani Cao

Department of Computer Science
City University of Hong Kong
Hong Kong, China
jncao2-c@my.cityu.edu.hk

Yang Liu*

Department of Computer Science
Florida State University
Tallahassee, USA
yl25r@fsu.edu

Kun Wang

Department of Computer Science
City University of Hong Kong
Hong Kong, China
kwang69-c@my.cityu.edu.hk

Zhenjiang Li*

Department of Computer Science
City University of Hong Kong
Hong Kong, China
zhenjiang.li@cityu.edu.hk

Abstract

Motor Imagery (MI) is an emerging Brain-Computer Interface (BCI) paradigm in which a person imagines a body movement without any physical action. By decoding the scalp-recorded electroencephalography (EEG) signals, BCIs can establish direct communication pathways to control external devices, offering significant potential in prosthetics, rehabilitation, and human-computer interaction. However, existing solutions remain difficult to deploy in practice. (i) Most employ independent, opaque models for each MI task. This fragmented methodology lacks a unified architectural foundation. Consequently, these models are trained in isolation and fail to learn robust representations from diverse datasets, which often results in modest performance. (ii) They primarily adopt fixed sensor deployment, whereas real-world setups vary in electrode number and placement, causing models trained on one configuration to fail under another. (iii) Performance degrades sharply under low-SNR conditions typical of consumer-grade EEG. Together, these limitations hinder the practical adoption of MI-based BCIs.

To address these challenges, we present NeuroPath, a well-designed neural architecture for robust MI decoding. NeuroPath takes high-level inspiration from the brain's signal pathway from cortex to scalp, utilizing a deep neural architecture with specialized modules for signal filtering, spatial representation learning, and feature classification, enabling unified decoding rather than task-specific black-box learning. To handle variations in electrode configurations, we introduce a spatially aware graph adapter that accommodates different electrode numbers and placements. To enhance robustness under low-SNR conditions, NeuroPath incorporates multimodal auxiliary training to refine EEG representations and stabilize performance on noisy, real-world data. Evaluations on three self-collected consumer-grade datasets and three public medical-grade datasets demonstrate that NeuroPath achieves superior performance.

*Corresponding authors.



This work is licensed under a Creative Commons Attribution 4.0 International License. *SenSys '26, Saint Malo, France*

© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2309-4/26/05
<https://doi.org/10.1145/3774906.3802770>

CCS Concepts

• **Human-centered computing** → **Ubiquitous and mobile computing systems and tools; Interaction techniques.**

Keywords

Brain-Computer Interface (BCI), motor imagery, mobile sensing

ACM Reference Format:

Jiani Cao, Kun Wang, Yang Liu, and Zhenjiang Li. 2026. NeuroPath: Practically Adopting Motor Imagery Decoding through EEG Signals. In *ACM/IEEE International Conference on Embedded Artificial Intelligence and Sensing Systems (SenSys '26)*, May 11–14, 2026, Saint Malo, France. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3774906.3802770>

1 Introduction

Brain-Computer Interface (BCI) technology [20] enables direct communication between the brain and external devices, unlocking applications in assistive control, rehabilitation, and human-computer interaction [27, 42, 61]. Among BCI paradigms, *Motor Imagery* (MI) is an intuitive, stimulus-independent approach: *electroencephalography (EEG) sensors record brain activity while a person imagines moving a body part from a first-person perspective, without actual movement* (Figure 1) [33, 45, 55]. Unlike other BCI paradigms such as Event-Related Potentials (ERPs) [23] or Steady-State Visually Evoked Potentials (SSVEPs) [22], which depend on external cues and can induce fatigue, MI aligns naturally with volitional control and offers a more intuitive means of interaction. Beyond practical interfaces, MI studies also deepen our understanding of the neural mechanisms underlying motor and cognitive functions [47].

A user's movement intention in MI arises from the synchronized firing of tens of thousands of neurons in motor cortical regions, generating neural source signals [59] with characteristic spatiotemporal patterns. As these signals travel toward the scalp, they traverse multiple biological layers, including the meninges, cerebrospinal fluid, skull, and scalp [11]. Each layer acts as a volume conductor and effective low-pass filter [31], causing strong attenuation, spatial mixing, and additional distortion from noise (see §2.1). As a result, the scalp-recorded EEG is a weak, distorted superposition of activities from multiple brain regions, which complicates

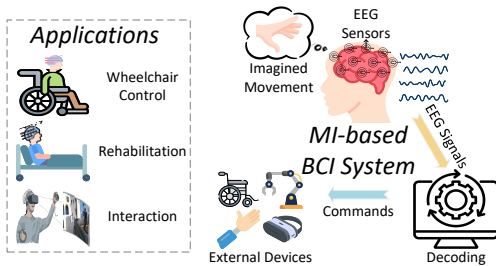


Figure 1: The process of MI-based BCI system and its potential applications.

accurate intent inference. Despite these complexities, recent approaches have harnessed machine learning to achieve encouraging results [6–8, 19, 21, 39, 68]. However, the wider adoption and practical deployment of the MI system remain difficult. Most existing approaches are designed as task-specific black-box models that lack a unified architectural foundation. This makes them fragile, struggling to aggregate knowledge from diverse datasets, adapt to variations in sensor configurations, or maintain performance in noisy, real-world conditions. Collectively, these limitations hinder the practical adoption of MI-based BCIs.

To fill this gap, we present **NeuroPath**, a new MI-based BCI framework that reliably and practically decodes imagined movements by addressing two core challenges:

1) The prevailing paradigm of developing isolated, task-specific black-box models. Most approaches employ independent, opaque models for each MI task. This fragmented, task-specific methodology lacks a unified architectural foundation. Consequently, these models are trained in isolation and fail to learn a robust and informative representation from the collective knowledge of diverse datasets. This not only leads to modest performance but also hinders interpretability, making it difficult to understand or trust the model’s decisions.

NeuroPath addresses this by introducing a well-designed neural architecture for robust MI decoding. The main idea is to use the brain’s forward signal pathway from cortex to scalp as a high-level design inspiration for our deep learning architecture, providing a structural rationale for our neural network components. Taking high-level design inspiration from the neural signal pathway [11, 58], NeuroPath organizes decoding into three coupled modules. First, a *signal filtering* stage mitigates noise and non-cortical interference. Second, a *spatial representation learning* stage extracts robust spatial features from scalp recordings. Third, a *feature aggregation and classification* stage processes these representations to infer the user’s imagined movement, providing unified and interpretable decoding rather than task-specific black-box mappings.

2) The inherent fragility of these models to electrode variability and signal noise. Recent designs primarily target the EEG devices or data available to them, while real-world configurations vary in electrode number and placement, causing the model design limited to a given electrode configuration and the model trained on one configuration to fail being used when deployed on another. Furthermore, when MI-based BCI systems are deployed with mobile EEG devices, performance may degrade sharply due to the low SNR of EEG signals from commercial EEG devices.

To address the electrode heterogeneity, we introduce a *spatially aware graph adapter* that projects inputs from arbitrary electrode configurations into a shared latent space. The adapter constructs a graph over electrodes based on their deployment coordinates (e.g., 10–10 system positions [2]) and learns layout-invariant embeddings that can gracefully handle missing channels, making the system to be applied to various EEG electrode configurations. To mitigate the low-SNR issue from consumer-grade mobile EEG, NeuroPath further employs *multimodal auxiliary training*. During training, auxiliary streams, such as automatically generated visual cues, serve as privileged information that refines EEG representations and promotes noise-robust feature learning. At inference time, however, NeuroPath relies solely on EEG inputs. Together, these two designs directly address the challenges of electrode variability and signal noise, ensuring robust performance across heterogeneous devices and real-world environments.

We integrate above designs into NeuroPath, implement a prototype with a consumer EEG device Emotiv FLEX 2 [3], and deploy it on a Google Pixel 7 smartphone. This mobile deployment serves as a crucial validation of its real-world practicality, demonstrating that NeuroPath operates with exceptional efficiency on resource-constrained hardware. To comprehensively evaluate its performance, we conduct extensive experiments on three self-collected datasets (NeuroPath-DS-32C, 16C and 8C, gathered from 12 volunteers) and three public benchmark datasets (BCIC-2a [12], BCIC-2b [35], and MI-KU [34]). We compare NeuroPath with the recent methods including LGL-BCI [44], CLTNet [21], and CIACNet [39]. Overall, NeuroPath consistently maintains competitive performance across all six datasets. We also examine the performance of NeuroPath under various settings, including different user hair lengths, device wearing positions, and ambient noise levels, etc. Results show that NeuroPath maintains robust and reliable performance in diverse settings.

In summary, this paper makes the following key contributions:

- We introduce NeuroPath, a deep learning architecture for MI decoding that takes high-level structural inspiration from the cortex-to-scalp signal pathway, employing three specialized neural modules (signal filtering, spatial representation learning, and feature classification) to yield robust decoding.
- We develop training mechanisms that make MI practical under real-world variability: a spatially aware graph adapter that handles heterogeneity in electrode settings (i.e., numbers and placements) by projecting EEG data into a unified latent space, and multimodal auxiliary training that uses auxiliary cues only during training to learn noise-robust EEG features while requiring EEG alone at inference.
- We develop NeuroPath and conduct extensive evaluations on both consumer-grade and medical-grade datasets. Results show that NeuroPath consistently maintains competitive performance across all datasets.

2 Preliminary

2.1 Basic Principles of Motor Imagery

MI refers to the mental rehearsal of a body movement from a first-person perspective without any actual physical execution. To decode this mental activity accurately, it is important to understand

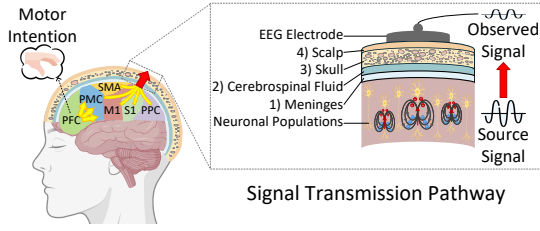


Figure 2: MI signal transmission pathway in the brain.

how the neural signals are generated, how they are sensed, and what characteristics they exhibit, as these very principles give rise to the core challenges in practical MI-BCI systems.

Generation and capture of neural signals. This process begins when a person imagines a specific movement (e.g., moving the left hand). Figure 2 illustrates the involved brain regions and the signal transmission pathway. The intention is first initiated in the Prefrontal Cortex (PFC), which then activates motor-related regions such as the Premotor Cortex (PMC), Supplementary Motor Area (SMA), and Primary Motor Cortex (M1). Neuronal populations in these regions fire *synchronously*, generating neural **source signals**. As these source signals propagate toward the scalp, they pass through several physiological layers including the 1) meninges, 2) cerebrospinal fluid, 3) skull, and 4) scalp, where they undergo strong low-pass filtering and volume conduction effects [31] before reaching the surface. These faint activities can be captured non-invasively using EEG, which records electrical potentials from electrodes placed on the scalp.

This complex signal journey from a deep cortical source to the scalp is why MI decoding is so challenging. The resulting scalp EEG is a weak, distorted superposition of activities from multiple brain regions. This inherent complexity explains why treating the decoding process as a simple black-box problem often yields fragile and non-interpretable models; such models fail to account for the underlying physics of signal mixing and distortion.

EEG rhythms & event-related (de)synchronization. EEG signals are commonly categorized into rhythms: δ - θ (1–7 Hz) linked to deep sleep and drowsiness, α (8–12 Hz) associated with MI and relaxation, β (13–30 Hz) related to active thinking and motor control, and γ (>30 Hz) reflecting higher cognitive functions. Among these, while the α rhythm is most prominently associated with MI, other rhythms also contribute valuable information since MI engages multiple cognitive and sensorimotor processes. For example, when a person imagines clenching the left hand, the power of the α rhythm decreases, a phenomenon known as *Event-Related Desynchronization (ERD)*. After the imagery task ends, the α power often rebounds, which is termed *Event-Related Synchronization (ERS)*. Similar ERD and ERS phenomena are observed in the β rhythm.

To quantify these changes, we use the **ERD or ERS index**, defined as the percentage change in band power relative to a baseline period [60], $ERD/ERS \text{ index} (\%) = \frac{A-R}{R} \times 100\%$, where A is the power during the task and R is the baseline power in the 8–30 Hz range (covering α and β rhythms). Negative values indicate ERD (power decrease), while positive values indicate ERS (power increase) (Figure 3(a)). These repeatable EEG patterns provide the basis for inferring a user’s intent, but they are often subtle and easily

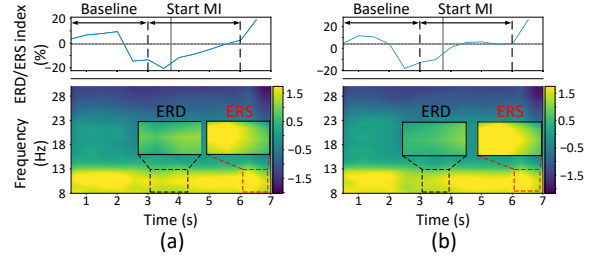


Figure 3: ERD and ERS patterns during motor imagery of clenching the (a) left hand and (b) right hand.

obscured by noise. In addition, imagery tasks for different limbs (e.g., imagining clenching the right hand versus the left hand) produce highly similar signal patterns (Figure 3(b), imagining clenching the right hand), making them nontrivial to distinguish.

2.2 Application Scenarios

Leveraging its non-invasive and intuitive nature, MI holds great potential in healthcare and daily interaction.

Limb replacement and rehabilitation. For patients with severe motor impairments, MI acts as a critical bridge between intention and action [38, 53]. It enables assistive control (e.g., wheelchair operation [52]) and allows clinicians to objectively monitor neural recovery even before physical function is regained [53].

Human-computer interaction. MI offers a silent, hands-free, and privacy-preserving control channel for the broader population [49]. It facilitates seamless interaction in smart homes [70] and provides an additional intuitive input modality for hands-busy scenarios like virtual reality and gaming [9].

2.3 System Overview

Figure 4 shows the overview of NeuroPath, which consists of two main components.

1) Brain-inspired MI decoder (§3). The core of NeuroPath is a decoder architecture inspired by the brain’s forward signal generation process. It draws structural inspiration from this pathway through three deep learning modules: signal filtering, spatial representation learning, and feature aggregation and classification. This design improves decoding accuracy, interpretability, and reliability, providing a unified architectural foundation for MI decoding.

2) Practical training framework (§4). To address the configuration heterogeneity and low-quality EEG signal issues, we further

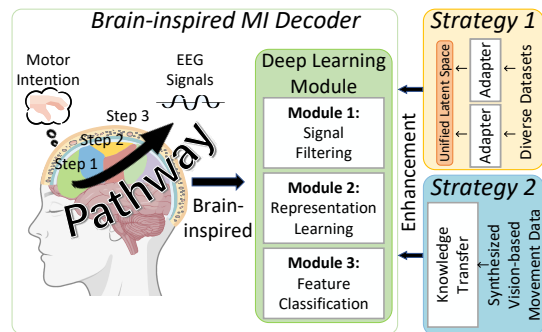


Figure 4: Overview of the NeuroPath design.

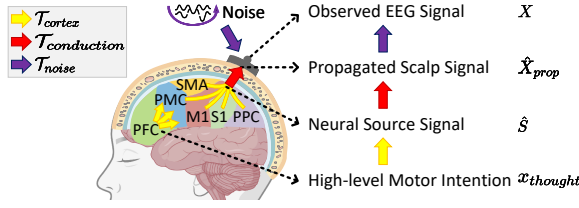


Figure 5: Illustration of the forward generation process.

propose two complementary training strategies that enhance the MI decoder proposed in §3. First, a spatially-aware graph adapter addresses heterogeneity in electrode settings (*i.e.*, varying numbers and placements) by projecting EEG data into a unified latent space. This makes the decoder agnostic to the input electrode configuration, achieved by leveraging multiple diverse datasets as robust priors during training. Second, a cross-modality knowledge transfer framework leverages synthesized vision-based movement data to guide training, enabling the decoder to learn more discriminative and robust motor features.

3 Brain-inspired MI Decoder

3.1 Design Rationale

The core idea of our decoder is to address the inherent complexity of how motor intentions are transformed into scalp-recorded EEG signals. A user’s motor intention, denoted as $x_{thought}$, is not directly observable on the scalp. Instead, it is first encoded by neuronal populations into neural source signals \hat{S} . These source signals then propagate through multiple biological layers, where they are attenuated and spatially mixed, resulting in the propagated scalp signal, denoted as \hat{X}_{prop} . Finally, this signal is contaminated by various external and physiological noise sources, producing the observed low-SNR EEG signal X :

$$x_{thought} \rightarrow \hat{S} \rightarrow \hat{X}_{prop} \rightarrow X.$$

Decoding thus becomes a challenging *ill-posed inverse problem*: how to accurately infer $x_{thought}$ from the noisy EEG X . While one could attempt to model and invert the entire biophysical process exactly, this is infeasible due to the incomplete understanding of brain mechanisms and the prohibitive computational cost.

Instead, we draw inspiration from the *encoder–decoder* paradigm in deep learning. The brain’s forward generation of signals (from $x_{thought}$ to X) can be regarded as a complex “encoder” governed by neurophysiological dynamics. Our idea is to construct a deep learning architecture inspired by these key stages, utilizing standard neural network components to process the data, rather than attempting a literal biological inversion. By using these physiological stages as high-level architectural priors, NeuroPath is guided toward more effective representation learning and robust decoding.

3.2 The Forward Generation Process

To guide the deep learning architecture design, we conceptually abstract the forward generation of EEG signals from $x_{thought}$ to X as a high-level composite function consisting of three stages (Figure 5):

$$X = [\mathcal{T}_{noise} \circ \mathcal{T}_{conduction} \circ \mathcal{T}_{cortex}](x_{thought}), \quad (1)$$

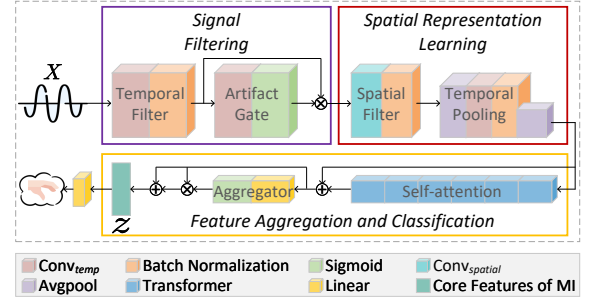


Figure 6: Brain-inspired MI decoder architecture.

where \circ denotes function composition.

1) Cortical processing transform \mathcal{T}_{cortex} :

$$\hat{S} = \mathcal{T}_{cortex}(x_{thought}). \quad (2)$$

This transform models the activity of motor-related brain regions such as the PMC, SMA, and M1 shown in Figure 5. It converts the high-level motor intention $x_{thought}$, initiated in the PFC, into multichannel neural source signals \hat{S} with characteristic spatiotemporal patterns [58]. These signals arise from the synchronized firing of neural ensembles in differential rhythmic patterns.

2) Volume conduction transform $\mathcal{T}_{conduction}$:

$$\hat{X}_{prop} = \mathcal{T}_{conduction}(\hat{S}). \quad (3)$$

As \hat{S} propagates outward, it passes through the meninges, cerebrospinal fluid, skull, and scalp. These biological layers act as low-pass filters and volume conductors [31], attenuating the signals and mixing them across electrodes. The result is \hat{X}_{prop} , a blurred linear superposition of multiple neural sources recorded at the scalp.

3) Noise contamination transform \mathcal{T}_{noise} :

$$X = \mathcal{T}_{noise}(\hat{X}_{prop}). \quad (4)$$

Finally, the scalp-level signals are contaminated by multiple sources of noise [36], including electromagnetic interference, device-related electronic noise, and physiological artifacts such as eye blinks or muscle activity. The observed EEG $X \in \mathbb{R}^{C \times T}$ is therefore a low-SNR signal, where C is the number of channels and T is the number of time samples.

3.3 Brain-inspired MI Decoder Architecture

Taking inspiration from the forward generation model, our MI decoder is designed with neural network modules that conceptually align with these physiological stages, mapping the noisy EEG sequence to a classification output. As shown in Figure 6, the decoding pipeline transforms the EEG input X into a compact feature vector z for classification. It consists of three modules, each drawing structural inspiration from a specific stage in the forward process.

1) Signal filtering module. Conceptually addressing the noise contamination stage, this module employs learnable temporal convolutions and gating mechanisms to mitigate noise and artifact contamination from the raw EEG X . It includes two sub-modules:

Learnable temporal filter. EEG signals contain both informative rhythms and a large amount of task-irrelevant activity across frequencies, making filtering essential to highlight MI-related bands

(e.g., α and β rhythms) while suppressing noise and unrelated components. Convolution is mathematically equivalent to finite impulse response (FIR) filtering, with the convolution kernel serving as the filter coefficients. Therefore, a temporal convolution can naturally function as a bank of parallel, learnable FIR filters, where larger kernels correspond to longer filter windows in the time domain. Instead of relying on manually defined passbands, the network learns optimal filter weights W_{temp} directly from data:

$$X_{filtered} = Conv_{temp}(X) = W_{temp} * X. \quad (5)$$

Adaptive artifact suppression gate. Even after temporal filtering, EEG often contains burst-like, non-stationary artifacts such as eye blinks, jaw clenching, or head movements. These artifacts typically appear with large amplitudes that can easily overshadow the subtle MI-related rhythms if not explicitly addressed. To mitigate this, we introduce an adaptive gating mechanism that learns to detect and suppress such artifacts. Specifically, a convolution followed by sigmoid activation generates a dynamic mask $G \in [0, 1]^{C \times T}$ that is synchronized with the signal:

$$G = \sigma(Conv_{gate}(X_{filtered})). \quad (6)$$

Values near 0 indicate corrupted segments to be down-weighted, while values near 1 preserve clean portions of the signal. This soft, element-wise weighting avoids hard signal removal and instead provides a fine-grained, data-driven suppression of artifacts.

The propagated scalp signal is then obtained through:

$$\hat{X}_{prop} = G \odot X_{filtered}, \quad (7)$$

where \odot denotes element-wise multiplication.

2) Spatial representation learning module. Conceptually addressing the volume conduction effect, this module uses spatial convolutional filters to learn robust, localized representations from the filtered scalp signals and extract key rhythmic features.

Spatial filter. Because each electrode records a superposition of activities from multiple cortical regions, direct use of scalp potentials makes it difficult to isolate MI-related signals. To address this, we apply a spatial convolution across all electrode channels that learns weights $W_{spatial}$ to optimally recombine the mixed signals. This operation acts as a data-driven spatial filter, similar in spirit to methods such as CSP (Common Spatial Patterns) [43], but optimized end-to-end within the network. By emphasizing discriminative spatial patterns, it separates overlapping sources and learns localized spatial representations \hat{S} :

$$\hat{S} = Conv_{spatial}(\hat{X}_{prop}) = W_{spatial} \cdot \hat{X}_{prop}. \quad (8)$$

Temporal pooling. Even after spatial convolution, the intermediate representations \hat{S} remain continuous time-series signals that need to be characterized in terms of their oscillatory content. MI is known to modulate specific neural rhythms (e.g., α and β), so extracting these rhythmic features is crucial for distinguishing different MI tasks. To this end, we apply temporal convolutions followed by average pooling, which together act as learnable band-pass filters and summarizers. The convolution learns frequency-selective patterns, while pooling aggregates them into compact feature tokens. The result is a set of discriminative rhythm representations F_{rhythm} :

$$F_{rhythm} = AvgPool(Conv_{rhythm}(\hat{S})). \quad (9)$$

Table 1: Comparison of key metrics between medical-grade and consumer-grade devices.

Parameter	Medical	Consumer
Electrode Type	Ag/AgCl Wet	Dry or Saline
Data Resolution	24-bit	16-bit
Voltage Sensitivity	0.39 nV	510 nV
Sampling Rate	High	Low
Channel Density	Dense	Sparse
Price (USD)	\$30,000+	\$500 - \$2,000

These features capture the essential oscillatory signatures of MI while reducing redundancy and noise, preparing them for high-level inference in the next stage.

3) Feature aggregation and classification module. Rather than processing raw signals, this final module performs classification based on the rhythmic features F_{rhythm} extracted by the previous layers.

Self-attention module. MI does not arise from a single brain region but from the coordinated activity of several sensorimotor areas working together. To correctly decode intention, it is therefore necessary to capture not only the strength of individual rhythms but also the interactions between them.

To achieve this, we use Transformer layers. A Transformer’s self-attention mechanism allows each feature token to dynamically weigh and relate to all others, regardless of their spatial or temporal distance. This ability to model long-range dependencies makes it well suited to represent the synergy across distributed brain regions. The output of the Transformer is then combined with the original rhythm features F_{rhythm} through a residual connection. This creates an enriched set of features $F_{enhanced}$ that benefits from both the local rhythmic patterns and their global interactions:

$$F_{enhanced} = F_{rhythm} + Transformer(F_{rhythm}). \quad (10)$$

Intention feature aggregator. Not all features contribute equally to distinguishing motor intentions. To emphasize the most informative components, we introduce a gating mechanism implemented with a linear layer followed by a sigmoid activation. The gate G_{agg} selectively weights the enhanced features through element-wise multiplication. The resulting sequence is then flattened into a compact vector z that captures the user’s motor intention.

$$G_{agg} = \sigma(Linear_{gate}(F_{enhanced})), \quad (11)$$

$$z = Flatten(G_{agg} \odot F_{enhanced} + F_{enhanced}). \quad (12)$$

Finally, the vector z is projected by a linear classifier to predict the motor intention $x_{thought}$. The decoder is trained end-to-end with a cross-entropy loss, encouraging the network to learn a feature space in which different motor intentions are clearly separable.

4 Practical Training Framework for Robust MI EEG Decoding

4.1 Design Motivation

The MI decoder proposed in §3 improves both accuracy and interpretability, providing a unified architectural foundation. However,

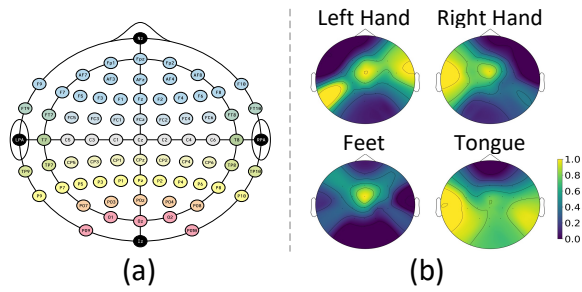


Figure 7: (a) International 10-10 electrode system. (b) Task-dependent importance of different channels, based on average signal amplitude in MI tasks.

for real-world deployment, we need to further address two practical challenges that hinder the widespread adoption of MI-BCIs.

First, EEG data is plagued by heterogeneity in electrode numbers and placements. This poses two critical problems. On one hand, it obstructs knowledge aggregation from diverse datasets, which are invaluable for training a powerful model but feature disparate electrode layouts. On the other hand, it creates a critical gap: a model trained for a specific electrode layout typically fails to work when deployed on a system with a different configuration. This fundamental mismatch between data sources, both across training sets and between training and deployment, severely limits a model’s scalability and real-world applicability.

Second, non-intrusive EEG recordings suffer from a low-SNR, a problem that is especially severe in consumer-grade devices. As shown in Table 1, compared with their high-cost, medical-grade counterparts, these consumer-grade devices typically use dry or saline electrodes, have lower resolution and sensitivity, and provide fewer recording channels. These limitations greatly weaken the captured MI-related rhythms: subtle neural oscillations may be attenuated below detection thresholds, while sparse spatial sampling blurs activity from distinct cortical regions. As a result, the signals collected by consumer devices have much lower SNR, making them significantly harder to decode reliably.

4.2 Addressing Electrode Heterogeneity

To address the challenge of electrode heterogeneity, we introduce a novel component designed to make our model agnostic to specific electrode layouts: the spatially-aware graph adapter. Its primary function is to transform EEG data from arbitrary layout into a unified, standardized representation. This approach directly solves the deployment mismatch problem and paves the way for aggregating knowledge from diverse data to further enhance model robustness.

1) Spatially-aware graph adapter. To make our model functional across diverse electrode configurations, we introduce a spatially-aware graph adapter in NeuroPath.

1.1) Insights. The spatially-aware graph adapter transforms heterogeneous EEG inputs of size (C, T) into a unified tensor (C_{fix}, T_{fix}) , enabling consistent processing by the shared MI decoder proposed in §3. In doing so, it not only makes pre-training on heterogeneous datasets feasible but also allows the learned features to transfer seamlessly to varied devices. Its design is guided by two insights:

Structured spatial prior. The international 10–10 system [2] provides a standardized map of electrode positions across the scalp, as shown in Figure 7(a). Different devices or datasets typically record from different subsets of these electrodes, leading to heterogeneity in channel numbers and placements. To handle this, we model electrodes as nodes in a graph, with edges encoding physical distances between locations. This representation preserves the true spatial layout of electrodes and enables consistent integration across datasets, rather than treating channels as an unordered list.

Dynamic channel importance. Even within the same electrode layout, not all electrodes contribute equally to every motor imagery task. For instance, hand motor imagery produces contralateral brain activity: imagining movement of the left hand primarily activates electrodes on the right side of the scalp, whereas imagining movement of the right hand primarily activates those on the left, as illustrated in Figure 7(b). Similar task-specific patterns can be observed for other imagery tasks such as feet or tongue movements. To capture this variability, we employ a graph attention mechanism that adaptively assigns weights to electrodes based on task relevance. This allows the model to focus on the most informative channels while down-weighting irrelevant ones.

1.2) Design. Building on the above insights, the adapter works in three steps for each heterogeneous dataset:

Step 1: Graph initialization. Each electrode is treated as a graph node. We assign initial node features based on signal statistics, including band power in δ - θ (1–7 Hz), α (8–12 Hz), β (13–30 Hz), and γ (31–45 Hz) rhythms, together with signal variance. At the same time, we construct the graph structure using the standardized 10–10 electrode coordinates: pairwise Euclidean distances between electrodes are converted into an adjacency matrix via a Gaussian kernel, which encodes spatial proximity as graph edges.

Step 2: Attention-based aggregation. The initialized graph and raw EEG signals are processed by a Graph Attention Network (GAT). The GAT learns to dynamically weight information from different electrodes, amplifying task-relevant channels while suppressing less informative ones. This produces an enriched feature representation reflecting both spatial structure and task relevance.

Step 3: Unified projection. Since different datasets have varying channel numbers and locations, the aggregated features are finally mapped to a fixed channel dimension C_{fix} (32 in our implementation) through a linear layer. The temporal dimension T is simultaneously aligned to T_{fix} using interpolation. This ensures all datasets are represented in a consistent tensor shape (C_{fix}, T_{fix}) for subsequent decoding.

The above steps transform heterogeneous EEG data (C, T) into a unified and informative representation (C_{fix}, T_{fix}) that preserves spatial structure and task-relevant information, ready for processing by the brain-inspired MI decoder.

2) Cross-dataset pre-training framework. With the proposed adapter, MI model can now process data from any electrode configuration in 10-10 system, resolving the critical challenge of electrode heterogeneity. This capability unlocks a powerful secondary benefit: we can now train a single, robust model by aggregating knowledge from multiple heterogeneous datasets simultaneously. To leverage this, we design a cross-dataset pre-training framework, illustrated in Figure 8, which enhances the model’s feature representations by

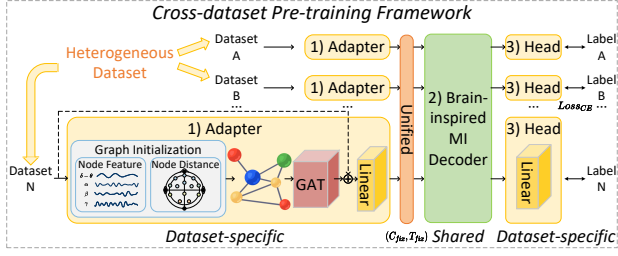


Figure 8: Cross-dataset pre-training framework integrating heterogeneous datasets through dataset-specific adapters and heads, with a shared MI decoder.

exposing it to a wide variety of data. The pipeline consists of three main components:

Dataset-specific adapters. Each heterogeneous dataset is first processed by its own adapter, which converts raw EEG into the unified representation (C_{fix}, T_{fix}) while preserving spatial structure and task-relevant information.

Shared MI decoder. The unified features are then passed into the brain-inspired MI decoder introduced in §3. This serves as a common backbone, enabling the model to learn shared neural feature priors across datasets.

Dataset-specific heads. Finally, the shared features are mapped to the label space of each dataset by a lightweight classification head, ensuring compatibility with different experimental tasks and label definitions.

Overall, this design offers a comprehensive solution to the challenge of electrode heterogeneity. First and foremost, the spatially-aware graph adapter solves the critical problem of deployment mismatch, making the model inherently flexible and scalable to new datasets and hardware configurations. Building on this foundation, the cross-dataset pre-training framework further enhances model robustness and performance. By learning from diverse data sources, the shared decoder acquires rich, transferable neural feature priors. We currently employ a sequential dataset iteration strategy. Specifically, the model iterates through the full content of each dataset sequentially within a single training epoch.

4.3 Addressing Low Signal Quality

To address the challenge of low signal quality, we introduce an auxiliary modality with clearer and more discriminative features to guide the fine-tuning phase after pre-training with §4.2. Rather than relying on additional sensors, this modality is generated through automated skeleton data synthesis. The key idea is that skeleton data provides a high-level structural representation of the instructed movement: it is inherently noise-free, requires minimal effort to generate with modern vision techniques, and captures distinct dynamic patterns for different imagined actions. As such, the skeleton modality serves as a low-cost but effective “teacher”, guiding the MI decoder (“student”) through cross-modality knowledge transfer.

1) Automated skeleton data generation. To obtain the skeleton modality without extra sensors, we leverage AIGC tools (e.g., vigo.ai [5]) to generate short videos of target MI actions. Given a frontal photo (optional) and a text prompt describing the action, the tool synthesizes the video, from which we extract 3D coordinates of 46 body, hand, and face keypoints using Google MediaPipe [4]

Finetune From Cross-dataset Pre-training Framework (§4.2)

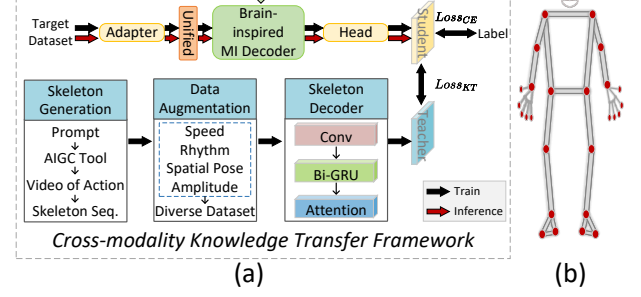


Figure 9: (a) Cross-modality knowledge transfer framework. (b) Skeleton keypoints extracted by MediaPipe.

(Figure 9(b)). This forms a high-quality skeleton sequence for each action. This incurs only a one-time, low-effort generation overhead. Because these skeleton sequences are generated from static text descriptions rather than recorded simultaneously with the user’s EEG, they share the same total duration but do not require strict frame-by-frame temporal alignment.

2) Data augmentation pipeline. Since one generated skeleton sequence is insufficient for robust training, we expand it into a diverse dataset by simulating natural variations in how people repeatedly perform the same action. This is achieved through four types of augmentations:

- **Speed variation:** the sequence is resampled at different rates to mimic actions performed faster or slower, reflecting natural differences in execution speed.
- **Rhythm variation:** we apply non-linear temporal warping so that different phases of the action (e.g., onset, peak, and offset) take variable durations, simulating irregular or personalized rhythms.
- **Spatial variation:** small random global rotations are applied to the 3D coordinates to account for posture changes such as leaning forward, which do not alter the core action.
- **Amplitude variation:** all coordinates are scaled by random factors to simulate changes in movement size, such as clenching more tightly or moving with smaller gestures.

3) Cross-modality knowledge transfer framework. We design a skeleton decoder as the teacher network (Figure 9(a)). It combines temporal convolutions for local motion patterns, a bidirectional GRU for long-range dependencies, and multi-head attention to emphasize key frames, producing compact and discriminative features. The MI decoder (student) is trained jointly with the skeleton decoder, using a composite objective:

$$Loss_{total} = Loss_{CE} + Loss_{KT},$$

where $Loss_{CE}$ is the cross-entropy loss with ground-truth labels, and $Loss_{KT}$ is the KL divergence that encourages the MI decoder to align its predictions with those of the skeleton decoder. This knowledge transfer allows the MI decoder to benefit from the clearer skeleton modality during training.

4) Inference. After training, only the MI decoder is retained for inference. It benefits from the transferred skeleton knowledge and

Table 2: Summary of datasets used for evaluation. The first three are self-collected under different channel configurations, while the last three are public benchmark datasets.

Dataset	Subjects	Channels	Device	Electrode Type	Resolution (Sensitivity)	Cost (USD)	Collection Time Per User
NeuroPath-DS-32C	12	32	Consumer-grade (Emotiv FLEX 2)	Saline-based	14 bit (510 nV)	~\$1,500	~40 min
NeuroPath-DS-16C	12	16					
NeuroPath-DS-8C	12	8					
BCIC-2a [12]	9	22	Medical-grade (e.g., g.tec)	Gel-based Ag/AgCl	High-precision (< 1 μ V)	> \$30,000	~58 min
BCIC-2b [35]	9	3					~70 min
MI-KU [34]	19	56	Medical-grade (BrainAmp)	Gel-based Ag/AgCl	24 bit (~0.4 nV)	> \$50,000	~47 min

achieves stronger decoding performance on low quality EEG signals, without requiring skeleton data at runtime.

5 Evaluation

5.1 Experimental Setup

Implementation. NeuroPath is built on a consumer-grade headset Emotiv FLEX 2 [3] as shown in Figure 10, and the model is implemented using Python 3.8 and PyTorch 1.12. The model training follows a two-stage procedure. We use the Adam optimizer for all training stages. First, we employ the cross-dataset pre-training framework (§4.2) to pre-train the shared MI decoder (§3) on heterogeneous public, medical-grade datasets (detailed below). This pre-training phase runs for 1000 epochs with a batch size of 16 and a learning rate of 0.001. After pre-training, the model weights, rich with prior knowledge, are saved. Subsequently, we use the cross-modality knowledge transfer framework (§4.3) to fine-tune the pre-trained model on the target consumer-grade, low-quality dataset. The training parameters for this stage are kept consistent with the pre-training phase, with 1000 epochs, a batch size of 16, and a learning rate of 0.001, respectively. All models are trained on an NVIDIA V100 GPU. After training, we deploy the model to a Google Pixel 7 phone using Pytorch Mobile for inference to validate its practical performance in a mobile environment.

Data collection. To comprehensively evaluate the performance of NeuroPath, we use six datasets, including three self-collected datasets and three benchmark public datasets as shown in Table 2.

1) NeuroPath-DS (Self-Collected). This dataset is collected using the Emotiv FLEX 2 [3], a consumer-grade device. This device supports flexible channel configurations up to 32 channels and has a relatively coarse signal resolution (*i.e.*, 510 nV) compared to medical-grade equipment (as low as 0.39 nV). To systematically evaluate the performance of NeuroPath under varying levels of practicality and to comprehensively investigate the trade-off between channel count and decoding accuracy, we construct three sub-datasets with different numbers of channels. These sub-datasets represent full-channel, balanced, and high-portability application scenarios. Their respective electrode layouts are shown in Figure 11(a), (b), and (c):

- **NeuroPath-DS-32C:** A full-channel configuration that includes all 32 channels supported by the device. It is designed

to serve as a performance benchmark and explore the system’s potential.

- **NeuroPath-DS-16C:** A balanced configuration with 16 channels. These channels are selected based on their relevance to motor imagery, prioritizing those covering the sensorimotor cortex to balance decoding performance and ease of wear.
- **NeuroPath-DS-8C:** A high-portability configuration with only 8 channels. The channel selection focuses on the most critical brain regions for MI (*e.g.*, near C3, C4, and Cz), representing a quick-setup, daily-use consumer scenario.

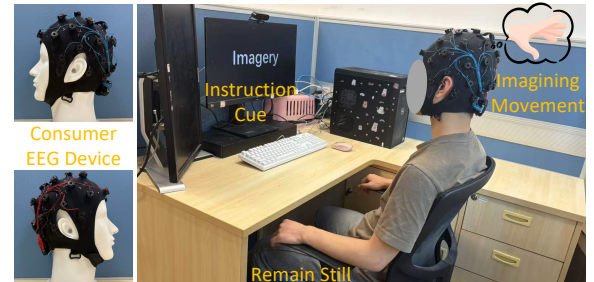


Figure 10: Illustration of the experimental setup.

Data for all sub-datasets are collected from twelve volunteers (four females and eight males), with ethical approval obtained from our institution. Before formal data collection, each volunteer performs the actual movements to activate the relevant motor cortex regions and to build an intuitive understanding of the imagery tasks. During data collection, users are explicitly instructed to keep their body physically still and perform mental rehearsal of the action. The procedure for one trial is shown in Figure 12. Volunteers sit comfortably facing a computer screen. At $t=0$ s, a fixation cross appears. At $t=2$ s, a prompt (*e.g.*, “clench left fist”, “clench right fist”, “raise both feet”, or “move tongue”) is displayed, and the order of prompts is randomized across trials. This visual instructional cue serves as the ground truth for the user’s motor intention in each trial. At $t=3$ s, the prompt disappears and the volunteer begins the corresponding motor imagery, which continues until $t=6$ s, when a “break” message indicates the trial’s end. We use the 4-second EEG segment from $t=2$ s to $t=6$ s as the effective imagery data for training and testing. The data are then split into training and testing sets at a 4:1 ratio to train a personalized model for each volunteer.

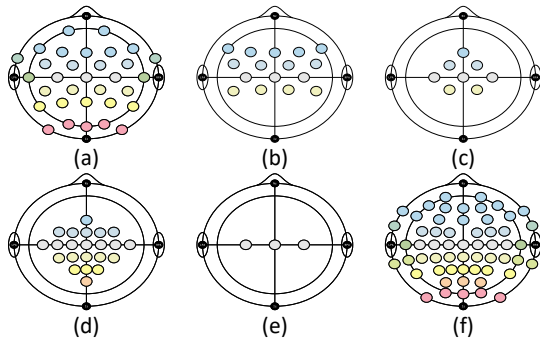


Figure 11: Different electrode numbers and placements for (a) NeuroPath-DS-32C, (b) NeuroPath-DS-16C, (c) NeuroPath-DS-8C, (d) BCIC-2a [12], (e) BCIC-2b [35] and (f) MI-KU [34].

2) **BCIC-2a** [12]. This public dataset contains EEG data from 9 subjects, recorded using 22 Ag/AgCl electrodes (a type of high-quality wet electrode) with the placement shown in Figure 11(d) at a sampling rate of 250 Hz. The paradigm is a four-class MI task involving the imagined movement of the left hand, right hand, both feet, and tongue. During a trial, a fixation cross appears at $t=0$ s. At $t=2$ s, a visual cue lasting 1.25 s instructs the subject to perform the imagery task, which continues until $t=6$ s. The cues for each class appear in a random order. We use the 4-second data from $t=2$ s to $t=6$ s as the effective imagery data. Each subject’s data includes a training and an evaluation session, each with 288 trials.

3) **BCIC-2b** [35]. This public dataset is also recorded from 9 subjects, using 3 Ag/AgCl electrodes (same high-quality type as in BCIC-2a) with the placement shown in Figure 11(e) at a 250 Hz sampling rate. The paradigm is a two-class MI task involving imagined movement of the left and right hand. The procedure is similar, with subjects performing a 4-second MI task after a visual cue. The cues for each class appear in a random order. Each subject’s data consists of five sessions, with each session containing 120 to 160 trials, split equally between the two classes. Data from the first three sessions constitute the training set, while data from the final two sessions form the evaluation set.

4) **MI-KU** [34]. This public dataset is collected using a BrainAmp amplifier, a device offering high-precision 24-bit resolution and excellent noise rejection, with 62 Ag/AgCl electrodes arranged as shown in Figure 11(f), at a sampling rate of 1000 Hz. As only 56 of its electrodes conform to the international 10-10 standard

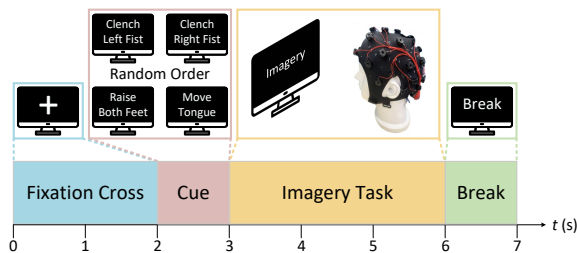


Figure 12: Data collection procedure of one trial in self-collected NeuroPath-DS. This process is similar to that of public medical-grade datasets, such as BCIC-2a [12].

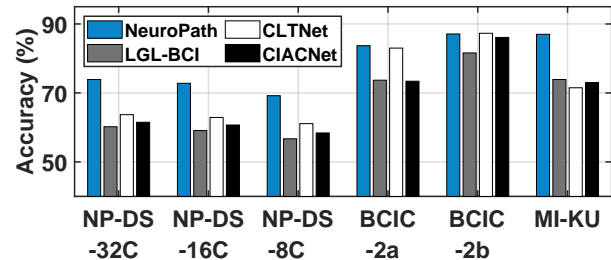


Figure 13: Overall performance. “NP” means “NeuroPath”.

system layout, we use only these 56 channels in our experiments. The dataset includes 54 subjects. The original paper notes that some subjects exhibit “MI BCI illiteracy”, meaning they cannot effectively produce distinguishable MI signals. To ensure the quality of our pre-training data and align with the paper’s findings, we select data from 19 subjects who are not “MI BCI illiterate” for our experiments. The paradigm is a two-class MI task involving the imagined movement of the left hand and right hand. Each trial begins with a 3-second preparation period, followed by a visual cue and a 4-second imagery task. The cues for each class appear in a random order. Each subject’s data includes two sessions, each containing a training phase and a testing phase, with 100 trials per phase (50 for each of the two classes).

Evaluation metric. We use accuracy as the evaluation metric. Accuracy is defined as the percentage of test trials correctly decoded out of the total number of test trials.

5.2 Overall Performance

We compare the following methods:

- **LGL-BCI** [44]. This is a recent end-to-end geometric learning approach designed for MI decoding.
- **CLTNet** [21]. This is a recent end-to-end hybrid deep learning architecture designed for MI decoding.
- **CIACNet** [39]. This is a recent end-to-end composite improved attention convolutional network designed for MI decoding.
- **NeuroPath.** The system proposed in this paper.

Figure 13 shows the overall performance of the four methods across six different datasets. Specifically, we first pre-train the shared MI decoder on the three medical-grade public datasets BCIC-2a, BCIC-2b, and MI-KU using a cross-dataset pre-training framework, which enables it to learn rich prior knowledge. Subsequently, we fine-tune the pre-trained model on each of the six target datasets. During the fine-tuning phase, we simultaneously activate the cross-modality knowledge transfer framework to further enhance the model’s decoding performance on specific tasks. Throughout the entire process, we strictly ensure that there is no overlap between the training and test sets.

We can see that NeuroPath consistently maintains competitive performance compared to the other three baselines on all six datasets with different electrode numbers and placements. To be specific, on the three challenging consumer-grade datasets (NeuroPath-DS with 32, 16, and 8 channels), NeuroPath’s decoding accuracies are 73.9%, 72.8%, and 69.2%, respectively. These results significantly

outperform LGL-BCI by margins of 12.5%–13.7%, CLTNet by 8.1%–10.2%, and CIACNet by 10.8%–12.4%. Furthermore, on the three medical-grade public datasets (BCIC-2a, BCIC-2b, and MI-KU), NeuroPath achieves decoding accuracies of 83.7%, 87.1%, and 87.0%, respectively, which outperform or are highly comparable to LGL-BCI (by 5.5%–13.1%), CLTNet (by -0.2%–15.5%), and CIACNet (by 1.0%–14.0%). Ultimately, unlike previous end-to-end methods, NeuroPath aims not only to achieve high decoding performance but also to address the critical challenges of practical deployment in heterogeneous electrode configurations and low-SNR scenarios.

5.3 Ablation Study.

We conduct an ablation study to systematically evaluate the effectiveness of each module in NeuroPath. To this end, we construct and evaluate the following intermediate versions of NeuroPath on all six datasets:

- **V1:** This version serves as the baseline, training the MI decoder (§3) end-to-end on each dataset independently.
- **V2-1:** Building on V1, this version applies the cross-dataset pre-training framework from §4.2. We first pre-train the shared MI decoder on the three medical-grade public datasets (BCIC-2a, BCIC-2b, and MI-KU) and then fine-tune it on each of the six target datasets.
- **V2-2:** Building on V1, this version introduces the cross-modality knowledge transfer framework from §4.3, which leverages generated skeleton data as a teacher signal to enhance the MI decoder’s performance.
- **V3 (Full):** This is the complete version of NeuroPath, integrating all proposed modules.

Table 3: Ablation study.

Dataset	V1	V2-1	V2-2	V3 (Full)
NeuroPath-DS-32C	67.0%	71.6%	69.3%	73.9%
NeuroPath-DS-16C	66.8%	69.3%	68.7%	72.8%
NeuroPath-DS-8C	66.1%	68.6%	68.5%	69.2%
BCIC-2a	78.1%	82.4%	80.1%	83.7%
BCIC-2b	85.8%	86.3%	85.9%	87.1%
MI-KU	78.4%	85.9%	81.3%	87.0%

As shown in Table 3, the baseline version (V1) achieves decoding accuracies of 67.0%, 66.8%, and 66.1% on the three consumer-grade NeuroPath-DS datasets, and 78.1%, 85.8%, and 78.4% on the three public datasets, respectively. With V2-1, the accuracies increase to 71.6%, 69.3%, and 68.6% on the NeuroPath-DS datasets, and to 82.4%, 86.3%, and 85.9% on the public datasets. This demonstrates that by pre-training on larger and higher-quality datasets, the model learns more robust neural feature priors. This knowledge is then effectively transferred to the target task during the fine-tuning phase, thereby enhancing decoding performance. Similarly, solely applying the cross-modality knowledge transfer strategy (V2-2) also yields comprehensive performance gains, boosting the accuracies to 69.3%, 68.7%, and 68.5% on the NeuroPath-DS datasets, and to 80.1%, 85.9%, and 81.3% on the public datasets. This proves that transferring knowledge from the clear and feature-rich skeleton modality can effectively guide the MI decoder in learning more discriminative motor representations from noisy EEG signals. Finally,

the complete version of NeuroPath (V3), which integrates all modules, achieves the best performance across all datasets, reaching accuracies of 73.9%, 72.8%, and 69.2% on the NeuroPath-DS datasets, and 83.7%, 87.1%, and 87.0% on the public datasets. This indicates a positive synergy between the two enhancement strategies, as their combination maximizes the model’s decoding capability.

5.4 Micro-benchmarks

We then conduct micro-benchmark experiments to evaluate how the performance of NeuroPath is impacted by various factors.

Different users. Figure 14(a) shows the performance of NeuroPath on the twelve different users in the NeuroPath-DS-32C dataset. It is evident that there are individual differences in decoding accuracy, ranging from 57.5% to 88.8%. This variance primarily stems from inherent differences in physiological structures and neural activity patterns among users, which is a common phenomenon in the BCI field [34]. This performance can be further enhanced by collecting more user-specific data, which is a practical approach for NeuroPath, as ground truth acquisition is straightforward and only requires the user to wear the device while following instructions.

Different hair lengths. We categorize users into three groups based on hair length: Short (< 5 cm), Medium (5-30 cm), and Long (> 30 cm). Figure 14(b) shows the average decoding performance for each group. The short-haired group achieves the highest accuracy at 79.8%, while the accuracies for the medium and long-haired groups decrease to 77.5% and 73.5%, respectively. This is because longer hair may form an insulating layer between electrodes and the scalp, which increases contact impedance and thereby degrades the quality of the EEG signal, affecting decoding performance.

Different head circumferences. We divide users into three groups based on head circumference: Small (male <56 cm, female <54 cm), Medium (male: [56 cm, 58 cm], female: [54 cm, 56 cm]), and Large (male >58 cm, female >56 cm). Figure 14(c) shows the average performance for each group. The medium circumference group performs best, reaching 80.0%, while the performance for the small and large circumference groups drops to 71.5% and 77.1%, respectively. This is likely because the fixed-size EEG cap we use provides the best fit for users with medium head circumferences. A head that is too small may result in poor electrode contact, while a head that is too large may cause discomfort from the overly tight cap, affecting signal quality. This issue can be addressed by providing a custom-fitted cap for each user.

Different wearing positions. In our experiments, we use the Cz electrode (at the center of the scalp) as the anchor for the default wearing position. We then shift the entire cap forward, backward, left, and right by 2 cm to simulate slight positional deviations that might occur in daily use. As shown in Figure 15(a), compared to the default position, performance slightly improves when the cap is worn backward but decreases when it is moved in other directions, with a leftward shift having the largest negative impact (-3.7%). Overall, the impact of minor positional shifts on performance is within an acceptable range, indicating that our design, particularly the graph attention adapter, can dynamically focus on important electrode information and has a degree of robustness to slight changes in position.

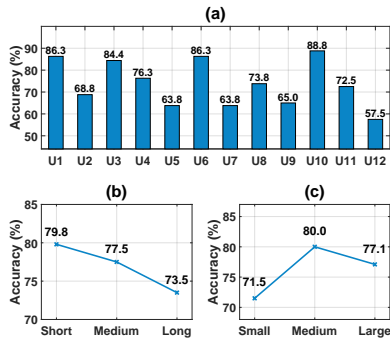


Figure 14: Performance across different (a) users, (b) hair lengths, and (c) head circumferences.

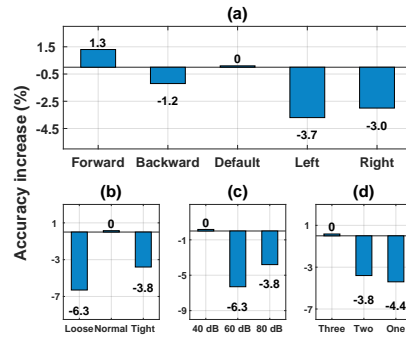


Figure 15: Performance under varying (a) wearing positions, (b) sensor attachments, (c) ambient volumes, and (d) pre-training data scales.

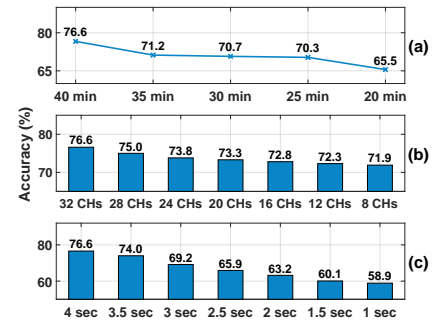


Figure 16: Performance with different (a) amounts of training data, (b) numbers of electrodes, and (c) trial data lengths.

Variations in strap tightness. We designate the state where users adjust the EEG cap to a comfortable level as “Normal”. Building on this, we loosen and tighten the straps by 1 cm, labeling these states as “Loose” and “Tight”, respectively. As shown in Figure 15(b), performance drops in the “Loose” state (-6.3%), as the poor fit leads to inadequate contact between the electrodes and the scalp, degrading signal quality. Performance also decreases in the “Tight” state (-3.8%), possibly because the excessive tightness causes user discomfort or muscle tension, which introduces additional noise.

Different ambient volumes. We evaluate the performance in different noise environments by simulating three scenarios: 40 dB (a quiet private bedroom), 60 dB (an office with people talking), and 80 dB (extreme weather with thunder outside). As shown in Figure 15(c), performance is best in the quiet environment (40 dB). As the ambient volume increases, performance declines, with accuracy decreasing by 6.3% and 3.8% in the 60 dB and 80 dB environments, respectively. This suggests that external noise, especially distracting conversations, can interfere with the user’s ability to concentrate on motor imagery, thereby affecting decoding performance.

Different amounts of pre-training data. We evaluate the impact of pre-training data scale by training on one, two, or all three public datasets. Compared to the three-dataset baseline, average decoding accuracy decreases by 3.8% and 4.4% when reducing the pre-training scale to two and one dataset, respectively (Figure 15(d)). This demonstrates that a larger and more diverse pre-training volume enables the shared decoder to learn more generalized and robust neural feature priors, significantly benefiting downstream performance.

Different amounts of training data. Since collecting EEG data is time-consuming and demanding for users, an ideal MI system must be data-efficient. To evaluate this, we examine the sensitivity of NeuroPath to the amount of training data by training the model with 40, 35, 30, 25, and 20 minutes of data per user. As shown in Figure 16(a), performance decreases as the training data are reduced. Notably, reducing the data from 40 minutes to 25 minutes (a 37.5% reduction) only lowers accuracy from 76.6% to 70.3%. Even with just 20 minutes of training data (a 50% reduction), the accuracy remains 65.5%. These results demonstrate that NeuroPath maintains good performance under limited data. Moreover, it suggest that with

more training data, the model has the potential to further improve its performance.

Different numbers of electrodes. The number of electrodes on a wearable EEG cap represents a critical trade-off between convenience and performance. We start with the default 32 electrodes and progressively reduce the number down to 8, prioritizing the retention of electrodes near the motor cortex area in each step. Figure 16(b) shows the decoding accuracy decreases gracefully as the number of electrodes is reduced. When the electrode count is lowered from 32 to 8, the accuracy drops from 76.6% to 71.9%, a performance loss that is within an acceptable range. This indicates that while a dense electrode layout provides richer information, NeuroPath can also operate effectively with fewer electrodes. This provides support for the future design of more portable, consumer-grade devices with fewer electrodes.

Different data lengths. In this experiment, we comprehensively evaluate the performance of NeuroPath with varying data lengths by reducing the trial length from the standard 4 seconds down to 1 second. As shown in Figure 16(c), when the trial length is reduced from 4 seconds to 1 second, the accuracy drops from 76.6% to 58.9%. This result indicates that NeuroPath maintains effective decoding capabilities when operating on certain shorter data segments (*e.g.*, greater than 2 seconds). However, when the window is further reduced (below 2 seconds), the accuracy experiences a more noticeable decline. This objectively illustrates the inherent trade-off between low latency and high decoding accuracy.

BCI illiterate population. In this experiment, we evaluate the performance of the remaining users in the MI-KU dataset, specifically the user group considered to be BCI illiterate. BCI illiteracy is a recognized physiological phenomenon in the field, referring to users who inherently cannot produce motor imagery EEG patterns that can be effectively detected by the system. We follow the general consensus in the field, specifically applying a decoding accuracy threshold below 75% to the subjects in the MI-KU dataset to define its BCI illiterate population. As shown in Table 4, the performance of this population (35 subjects in total) ranges from 54.0% to 74.0%, with an average accuracy of 62.2%. This physiological limitation poses a concrete impact on real-world deployment, highlighting the boundary of current MI-BCI applicability.

Table 4: Performance on the BCI illiterate population.

ID	Acc(%)	ID	Acc(%)	ID	Acc(%)	ID	Acc(%)	ID	Acc(%)
S01	62.5	S05	64.0	S07	65.5	S08	65.0	S09	57.5
S10	57.0	S11	66.0	S12	68.0	S13	61.5	S14	56.5
S15	66.5	S17	54.5	S19	63.0	S20	71.5	S22	57.5
S23	57.0	S24	60.5	S28	63.0	S31	59.5	S32	71.0
S34	54.0	S37	73.0	S38	59.0	S39	55.0	S40	64.0
S41	59.0	S45	74.0	S46	74.0	S48	63.0	S49	55.5
S50	54.0	S51	57.5	S52	68.5	S53	57.0	S54	61.0

5.5 System Overhead

To evaluate real-world practicality, we deploy NeuroPath on a Google Pixel 7 (Android 13, 4355 mAh battery) via PyTorch Mobile. Table 5 summarizes the system overhead metrics.

Table 5: System overhead of NeuroPath.

Metric	Value
Model Parameters	60,176
Computational Complexity	352.17 M-Adds
Inference Latency (per trial)	21.22 ms
Power Consumption	145.10 mAh
Battery Impact	3.35% of total battery

1) Computational cost. The model in NeuroPath is exceptionally lightweight, requiring only 60.2 K parameters and 352.17 M-Adds per inference. This load is easily manageable by standard mobile CPUs without specialized hardware accelerators. **2) Real-time performance.** NeuroPath processes a complete 4-second MI trial in just 21.22 ms (averaged over 100 runs). This latency constitutes less than 0.6% of the data window’s duration (21.22 ms / 4000 ms), ensuring that the system can provide immediate feedback without any perceptible delay to the user. **3) Energy efficiency.** Measured via Android’s Batterystats and Battery Historian [1], the power consumption is 145.10 mAh (3.35% of total capacity). This low footprint supports nearly 30 hours of continuous operation on a single charge, confirming its viability for daily use.

6 Discussion

Physiological limitations and BCI illiteracy. A fundamental challenge for general deployment is BCI illiteracy, which refers to the inherent inability of certain users to produce distinct and detectable EEG patterns during motor imagery. Currently, NeuroPath cannot resolve this physiological limitation. Future solutions may require significantly extending user BCI training to stimulate neural responses, or integrating alternative control channels (*e.g.*, eye-tracking or electromyography) to ensure universal accessibility.

Robustness against physical motion. A current limitation of NeuroPath is its inability to handle strong physical motion, as muscle-generated electrical signals easily overpower subtle MI neural oscillations. However, since the MI paradigm primarily targets individuals with severe motor impairments, scenarios involving active physical movement fall largely outside its intended scope.

Scalability to complex control. The number of classes in current evaluation is relatively small. However, this setup is consistent with mainstream protocols in the motor imagery field and largely addresses the core interaction needs of typical BCI applications, such

as wheelchair control and discrete smart home interactions. In the future, with hardware upgrades and improved signal resolution, it is possible to extend this framework to more complex, fine-grained, or continuous control scenarios.

7 Related Work

We review the related work of this paper in this section.

Brain-Computer Interface (BCI) paradigms. Non-invasive EEG-based BCIs are highly portable and suitable for consumer use [18, 50, 64]. While stimulus-driven paradigms like SVEP [22] and P300 [23] cause user fatigue [10], Motor Imagery (MI) provides an intuitive, internally-driven control channel [45, 55]. However, widespread MI-BCI deployment is hindered by complex signal generation and the inherently low SNR of consumer-grade EEG [11].

Motor Imagery (MI) decoding methods. Early MI decoding relied on feature engineering (*e.g.*, CSP [30]), while recent deep learning models directly learn spatiotemporal features [6, 32, 54]. Despite high accuracy, existing models face critical limitations: (1) they lack neurophysiological interpretability [44, 68]; (2) they are typically trained on medical-grade datasets and fail to generalize to the low-quality signals of consumer devices [69]; and (3) they assume fixed electrode layouts [7, 21, 39]. To address these gaps, NeuroPath proposes a brain-inspired architecture paired with cross-modality and cross-dataset training strategies, ensuring robustness across heterogeneous, low-quality datasets.

Practicality of mobile sensing systems. Enhancing consumer-grade hardware practicality is central to mobile sensing [14, 17, 29, 65]. Moreover, the broader ecosystem of these mobile systems is continuously enriched by system-level advances in edge inference [56, 57, 62, 63] and scalable low-power wireless communications [24, 25, 66]. Prior work extracts robust features from weak signals using COTS devices [13, 16, 26], advanced signal processing [37, 40, 67], physical modeling [15, 46], and robust learning strategies [28, 41, 48, 51]. Inspired by these works, we tackle the unsolved challenge of robust decoding on low-quality consumer EEG by utilizing neurophysiology as a high-level architectural inspiration and by developing training strategies that generalize across heterogeneous datasets in the design of NeuroPath in this paper.

8 Conclusion

In this paper, we present NeuroPath, a well-designed neural architecture addressing practical Motor Imagery (MI) BCI challenges. Departing from conventional fragmented models, it introduces a unified deep learning framework inspired by the brain’s signal generation pathway. A spatially aware graph adapter handles variable sensor configurations, while a multimodal auxiliary training strategy enhances robustness against low-SNR consumer-grade devices. Extensive evaluations across three public medical-grade and three self-collected consumer-grade datasets demonstrate that NeuroPath consistently outperforms recent methods.

Acknowledgments

Zhenjiang Li was supported by the GRF grants from Research Grants Council of Hong Kong (CityU 11205624 and CityU 11213622). Corresponding authors: Yang Liu and Zhenjiang Li.

References

- [1] [n. d.]. Battery Historian. <https://developer.android.com/topic/performance/power/setup-battery-historian/>.
- [2] [n. d.]. EEG 10-10 system. https://commons.wikimedia.org/wiki/File:EEG-10-10_system_with_additional_information.svg/.
- [3] [n. d.]. Emotiv products. https://www.emotiv.com/products/flex-saline?srsId=AfmBOopC0wju9qU8U_cPRGRJkVSQ07aduezXB4eKQkwS8ZhJqy1bv5CR/.
- [4] [n. d.]. Google MediaPipe. <https://ai.google.dev/edge/mediapipe/solutions/guide/>.
- [5] [n. d.]. vivago.ai. <https://vivago.ai/home/>.
- [6] Hamdi Altaheri, Ghulam Muhammad, and Mansour Alsulaiman. 2022. Physics-informed attention temporal convolutional network for EEG-based motor imagery classification. *IEEE transactions on industrial informatics* (2022).
- [7] Hamdi Altaheri, Ghulam Muhammad, Mansour Alsulaiman, Syed Umar Amin, Ghadir Ali Altuwajiri, Wadood Abdul, Mohamed A Bencherif, and Mohammed Faisal. 2023. Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: A review. *Neural Computing and Applications* (2023).
- [8] Sion An, Soopil Kim, Philip Chikontwe, and Sang Hyun Park. 2023. Dual attention relation network with fine-tuning for few-shot EEG motor imagery classification. *IEEE Transactions on Neural Networks and Learning Systems* (2023).
- [9] Pasquale Arpaia, Antonio Esposito, Enza Galasso, Fortuna Galdieri, Angela Natalizio, Marco Parvis, Michael Sommeling, and Mattia Volpe. 2024. Endless runner game in virtual reality controlled by a self-paced brain-computer interface based on EEG and motor imagery. In *International Conference on Extended Reality*.
- [10] Maedeh Azadi Moghadam and Ali Maleki. 2023. Fatigue factors and fatigue indices in SSVEP-based brain-computer interfaces: a systematic review and meta-analysis. *Frontiers in Human Neuroscience* (2023).
- [11] Mark Bear, Barry Connors, and Michael A Paradiso. 2020. *Neuroscience: Exploring the brain, enhanced edition: Exploring the brain*.
- [12] Clemens Brunner, Robert Leeb, Gernot Müller-Putz, Alois Schlögl, and Gert Pfurtscheller. 2008. BCI Competition 2008–Graz data set A. *Institute for knowledge discovery, Graz University of Technology, Austria* (2008).
- [13] Kayla-Jade Butkow, Ting Dang, Andrea Ferlini, Dong Ma, Yang Liu, and Cecilia Mascolo. 2024. An evaluation of heart rate monitoring with in-ear microphones under motion. *Pervasive and Mobile Computing* (2024).
- [14] Jiani Cao, Jiesong Chen, Chengdong Lin, Yang Liu, Kun Wang, and Zhenjiang Li. 2024. Practical gaze tracking on any surface with your phone. *IEEE Transactions on Mobile Computing* (2024).
- [15] Jiani Cao, Chengdong Lin, Yang Liu, and Zhenjiang Li. 2022. Gaze tracking on any surface with your phone. In *Proc. of ACM SenSys*.
- [16] Jiani Cao, Yang Liu, Lixiang Han, and Zhenjiang Li. 2024. Finger tracking using wrist-worn EMG sensors. *IEEE Transactions on Mobile Computing* (2024).
- [17] Tao Chen, Yongjie Yang, Xiaoran Fan, Xiuzhen Guo, Jie Xiong, and Longfei Shanguan. 2024. Exploring the feasibility of remote cardiac auscultation using earphones. In *Proc. of ACM MobiCom*.
- [18] Alexandre Défossez, Charlotte Caucheteux, Jérémy Rapin, Ori Kabeli, and Jean-Rémi King. 2023. Decoding speech perception from non-invasive brain recordings. *Nature Machine Intelligence* (2023).
- [19] Yidan Ding, Chalisa Udompanyawit, Yisha Zhang, and Bin He. 2025. EEG-based brain-computer interface enables real-time robotic hand control at individual finger level. *Nature Communications* 16, 1 (2025), 1–20.
- [20] Bradley J Edelman, Shuailei Zhang, Gerwin Schalk, Peter Brunner, Gernot Müller-Putz, Cuntai Guan, and Bin He. 2024. Non-invasive brain-computer interfaces: state of the art and trends. *IEEE reviews in biomedical engineering* (2024).
- [21] He Gu, Tingwei Chen, Xiao Ma, Mengyuan Zhang, Yan Sun, and Jian Zhao. 2025. Cltnet: a hybrid deep learning model for motor imagery classification. *Brain Sciences* (2025).
- [22] Ning Guo, Xiaojun Wang, Dehao Duanmu, Xin Huang, Xiaodong Li, Yunli Fan, Hailan Li, Yongquan Liu, Eric Hui Kwong Yeung, Michael Kai Tsun To, et al. 2022. SSVEP-based brain computer interface controlled soft robotic glove for post-stroke hand function rehabilitation. *IEEE transactions on neural systems and rehabilitation engineering* (2022).
- [23] Christoph S Herrmann and Robert T Knight. 2001. Mechanisms of human attention: event-related potentials and oscillations. *Neuroscience & Biobehavioral Reviews* (2001).
- [24] Ningning Hou, Yifeng Wang, Xianjin Xia, Shiming Yu, Yuanqing Zheng, and Tao Gu. 2025. Molora: Intelligent mobile antenna system for enhanced lora reception in urban environments. In *Proc. of ACM SenSys*.
- [25] Ningning Hou, Xianjin Xia, Yifeng Wang, and Yuanqing Zheng. 2024. One shot for all: Quick and accurate data aggregation for LPWANS. *IEEE/ACM Transactions on Networking* (2024).
- [26] Changshuo Hu, Thivya Kandappu, Yang Liu, Cecilia Mascolo, and Dong Ma. 2024. BreathPro: Monitoring breathing mode during running with earables. *Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2024).
- [27] Md Kafiul Islam and Amir Rastegarnia. 2023. Recent advances in EEG (non-invasive) based BCI applications. *Frontiers in Computational Neuroscience* (2023).
- [28] Sijie Ji, Yaxiong Xie, and Mo Li. 2022. SiFall: Practical online fall detection with RF sensing. In *Proc. of ACM SenSys*.
- [29] Sijie Ji, Xuanye Zhang, Yuanqing Zheng, and Mo Li. 2023. Construct 3d hand skeleton with commercial wifi. In *Proc. of ACM SenSys*.
- [30] Aimin Jiang, Jing Shang, Xiaofeng Liu, Yibin Tang, Hon Keung Kwan, and Yanping Zhu. 2020. Efficient CSP algorithm with spatio-temporal filtering for motor imagery classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (2020).
- [31] Eric R Kandel, James H Schwartz, Thomas M Jessell, Steven Siegelbaum, A James Hudspeth, Sarah Mack, et al. 2000. *Principles of neural science*.
- [32] Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, Chou P Hung, and Brent J Lance. 2018. EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces. *Journal of neural engineering* (2018).
- [33] Florent Lebon, Christian Collet, and Aymeric Guillot. 2010. Benefits of motor imagery training on muscle strength. *The Journal of Strength & Conditioning Research* (2010).
- [34] Min-Ho Lee, O-Yeon Kwon, Yong-Jeong Kim, Hong-Kyung Kim, Young-Eun Lee, John Williamson, Siamac Fazli, and Seong-Whan Lee. 2019. EEG dataset and OpenBMI toolbox for three BCI paradigms: An investigation into BCI illiteracy. *GigaScience* (2019).
- [35] Robert Leeb, Clemens Brunner, G Müller-Putz, A Schlögl, and GJGUOT Pfurtscheller. 2008. BCI Competition 2008–Graz data set B. *Institute for knowledge discovery, Graz University of Technology, Austria* (2008).
- [36] Bohao Li, Tianshuo Cheng, and Zexuan Guo. 2021. A review of EEG acquisition, processing and application. In *Journal of Physics: Conference Series*.
- [37] Xin Li, Yinghui He, and Jun Luo. 2025. μ Ceiver-Fi: Exploiting Spectrum Resources of Multi-link Receiver for Fine-Granularity Wi-Fi Sensing. In *Proc. of ACM MobiCom*.
- [38] Wenzhe Liao, Jiahao Li, Xuesong Zhang, and Chen Li. 2023. Motor imagery brain-computer interface rehabilitation system enhances upper limb performance and improves brain activity in stroke patients: A clinical study. *Frontiers in Human Neuroscience* (2023).
- [39] Wenzhe Liao, Zipeng Miao, Shuaibo Liang, Linyan Zhang, and Chen Li. 2025. A composite improved attention convolutional network for motor imagery EEG classification. *Frontiers in Neuroscience* (2025).
- [40] Yang Liu, Kayla-Jade Butkow, Jake Stuchbury-Wass, Adam Pullin, Dong Ma, and Cecilia Mascolo. 2025. Respear: Earable-based robust respiratory rate monitoring. In *IEEE PerCom*.
- [41] Yang Liu, Qiang Yang, Kayla-Jade Butkow, Jake Stuchbury-Wass, Dong Ma, and Cecilia Mascolo. 2025. EarMeter: Continuous Respiration Volume Monitoring with Earables. *Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2025).
- [42] Fabien Lotte, Josef Faller, Christoph Guger, Yann Renard, Gert Pfurtscheller, Anatole Lécuyer, and Robert Leeb. 2012. Combining BCI with virtual reality: towards new applications and improved BCI. In *Towards practical brain-computer interfaces: Bridging the gap from research to real-world applications*.
- [43] Fabien Lotte and Cuntai Guan. 2010. Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms. *IEEE Transactions on biomedical Engineering* (2010).
- [44] Jianchao Lu, Yuzhe Tian, Yang Zhang, Quan Z Sheng, and Xi Zheng. 2025. LGL-BCI: A Motor-Imagery-Based Brain-Computer Interface with Geometric Learning. *Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2025).
- [45] Silvia Marchesotti, Michela Bassolino, Andrea Serino, Hannes Bleuler, and Olaf Blanke. 2016. Quantifying the role of motor imagery in brain-machine interfaces. *Scientific reports* (2016).
- [46] Nishant Mehrotra, Divyanshu Pandey, Akarsh Prabhakara, Yawen Liu, Swarun Kumar, and Ashutosh Sabharwal. 2024. Hydra: Exploiting multi-bounce scattering for beyond-field-of-view mmwave radar. In *Proc. of ACM MobiCom*.
- [47] Jörn Munzert, Britta Lorey, and Karen Zentgraf. 2009. Cognitive motor processes: the role of motor imagery in the study of motor representations. *Brain research reviews* (2009).
- [48] Jingping Nie, Yuang Fan, Ziyi Xuan, Minghui Zhao, Runxi Wan, Matthias Preindl, and Xiaofan Jiang. 2025. SoundTrack: A Contactless Mobile Solution for Real-time Running Metric Estimation for Treadmill Running in the Wild. *Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2025).
- [49] Chukwuemeka Nwagu, Alaa AlSlaity, and Rita Orji. 2023. EEG-based brain-computer interactions in immersive virtual and augmented reality: A systematic review. *Proc. of the ACM on Human-Computer Interaction* (2023).
- [50] Jinhui Ouyang, Mingzhu Wu, Xinglin Li, Hanhui Deng, Zhanpeng Jin, and Di Wu. 2024. NeuroBCI: multi-brain to multi-robot interaction through EEG-adaptive neural networks and semantic communications. *IEEE Transactions on Mobile Computing* (2024).
- [51] Xiaomin Ouyang, Xian Shuai, Yang Li, Li Pan, Xifan Zhang, Heming Fu, Sitong Cheng, Xinyan Wang, Shihua Cao, Jiang Xin, et al. 2024. ADMarker: A Multi-Modal Federated Learning System for Monitoring Digital Biomarkers of Alzheimer's Disease. In *Proc. of ACM MobiCom*.

- [52] Arrigo Palumbo, Vera Gramigna, Barbara Calabrese, and Nicola Ielpo. 2021. Motor-imagery EEG-based BCIs in wheelchair movement and control: A systematic literature review. *Sensors* (2021).
- [53] Elodie Saruco, Arnaud Saimpont, Franck Di Rienzo, Benjamin De Witte, Isabelle Laroyenne, Fanny Matéo, Marion Lapenderie, Sarah Goutte Solard, Isabelle Perretant, Charlotte Frenot, et al. 2024. Towards efficient motor imagery interventions after lower-limb amputation. *Journal of NeuroEngineering and Rehabilitation* (2024).
- [54] Robin Tibor Schirrmeyer, Jost Tobias Springenberg, Lukas Dominique Josef Fiederer, Martin Glasstetter, Katharina Eggensperger, Michael Tangermann, Frank Hutter, Wolfram Burgard, and Tonio Ball. 2017. Deep learning with convolutional neural networks for EEG decoding and visualization. *Human brain mapping* (2017).
- [55] Corina Schuster, Roger Hilfiker, Oliver Amft, Anne Scheidhauer, Brian Andrews, Jenny Butler, Udo Kischka, and Thierry Ettl. 2011. Best practice for motor imagery: a systematic literature review on motor imagery training elements in five different disciplines. *BMC medicine* (2011).
- [56] Leming Shen, Qiang Yang, Kaiyan Cui, Yuanqing Zheng, Xiao-Yong Wei, Jianwei Liu, and Jinsong Han. 2024. Fedconv: A learning-on-model paradigm for heterogeneous federated clients. In *Proc. of ACM MobiSys*.
- [57] Leming Shen, Qiang Yang, Xinyu Huang, Zijiang Ma, and Yuanqing Zheng. 2025. Gpiot: Tailoring small language models for iot program synthesis and development. In *Proc. of ACM SenSys*.
- [58] Ana Solodkin, Petr Hlustik, E Elinor Chen, and Steven L Small. 2004. Fine modulation in network activation during motor execution and motor imagery. *Cerebral cortex* (2004).
- [59] Wing-kin Tam, Tong Wu, Qi Zhao, Edward Keefer, and Zhi Yang. 2019. Human motor decoding from neural signals: a review. *BMC Biomedical Engineering* (2019).
- [60] Zhichuan Tang, Shouqian Sun, Sanyuan Zhang, Yumiao Chen, Chao Li, and Shi Chen. 2016. A brain-machine interface based on ERD/ERS for an upper-limb exoskeleton control. *Sensors* (2016).
- [61] Kaido Värbu, Naveed Muhammad, and Yar Muhammad. 2022. Past, present, and future of EEG-based BCI applications. *Sensors* (2022).
- [62] Kun Wang, Jiani Cao, Zimu Zhou, and Zhenjiang Li. 2024. Swapnet: Efficient swapping for dnn inference on edge ai devices beyond the memory budget. *IEEE Transactions on Mobile Computing* (2024).
- [63] Kun Wang, Zimu Zhou, and Zhenjiang Li. 2024. LATTE: Layer algorithm-aware training time estimation for heterogeneous federated learning. In *Proc. of ACM MobiCom*.
- [64] Di Wu, Jinhui Ouyang, Ningyi Dai, Mingzhu Wu, Haodan Tan, Hanhui Deng, Yongmei Fan, Dakuo Wang, and Zhanpeng Jin. 2022. DeepBrain: Enabling fine-grained brain-robot interaction through human-centered learning of coarse EEG signals from low-cost devices. *Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2022).
- [65] Qiang Yang, Yang Liu, Jake Stuchbury-Wass, Mathias Ciliberto, Tobias Röddiger, Kayla-Jade Butkow, Adam Luke Pullin, Emeli Panariti, Dong Ma, and Cecilia Mascolo. 2025. HearForce: Force Estimation for Manual Toothbrushing with Earables. *Proc. of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2025).
- [66] Shiming Yu, Xianjin Xia, Ningning Hou, Yuanqing Zheng, and Tao Gu. 2024. Revolutionizing lora gateway with xgate: Scalable concurrent transmission across massive logical channels. In *Proc. of ACM MobiCom*.
- [67] Jonghyuk Yun, Kyoosik Lee, Kichang Lee, Bangjie Sun, Jaeho Jeon, Jeonggil Ko, Inseok Hwang, and Jun Han. 2024. PowDew: Detecting Counterfeit Powdered Food Products using a Commodity Smartphone. In *Proc. of ACM MobiSys*.
- [68] Xiang Zhang, Lina Yao, Quan Z Sheng, Salil S Kanhere, Tao Gu, and Dalin Zhang. 2018. Converting your thoughts to texts: Enabling brain typing via deep feature learning of eeg signals. In *IEEE PerCom*.
- [69] Wei Zhao, Xiaolu Jiang, Baocan Zhang, Shixiao Xiao, and Sujun Weng. 2024. CTNet: a convolutional transformer network for EEG-based motor imagery classification. *Scientific reports* (2024).
- [70] Wei Zhuang, Yixian Shen, Lu Li, Chunming Gao, and Dong Dai. 2020. A brain-computer interface system for smart home control based on single trial motor imagery EEG. *International Journal of Sensor Networks* (2020).