IJRCS
International Journal of Robotics and Control Systems
Vol. 6, No. 2, 2026, pp. 1139-1164
ISSN 2775-2658
http://pubs2.ascee.org/index.php/ijrcs
ASCEE

Spatiotemporal Analysis of Time Window Length in Multi-Directional Motor Imagery Classification Using Emotiv Insight

Yamin Thwe [a,1], Dechrit Maneetham [a,2,*], Padma Nyoman Crisnapati [a,3], Myo Min Aung [a,4]

[a] Department of Mechatronics Engineering, Rajamangala University of Technology Thanyaburi (RMUTT), Pathum Thani 12120, Thailand
[1] yamin_t@mail.rmutt.ac.th; [2] dechrit_m@rmutt.ac.th; [3] crisnapati@rmutt.ac.th; [4] myomin_a@rmutt.ac.th
* Corresponding Author

ARTICLE INFO

Article history
Received January 04, 2026
Revised March 01, 2026
Accepted April 13, 2026

Keywords
Biomedical Engineering;
Motor Imagery;
Brain Computer Interface;
Emotive Insight;
Consumer-Grade

ABSTRACT

This paper proposes to examine the feasibility of classifying multi-directional motor imagery tasks by using the low-channel electroencephalography signal with a consumer-grade headset, thus trying to fulfill the demand for more practical and affordable brain-computer interfaces. EEG signals were recorded from 53 participants under a video-guided MI paradigm, with preprocessing including channel selection, average referencing, band-pass filtering, ICA-based signal inspection, and temporal windowing to construct sequential inputs for deep learning models, while participants maintained central gaze fixation throughout all tasks to minimize potential directional eye-movement confounds. Signal quality was assessed using independent component analysis prior to epoch extraction. The proposed method is tested with a convolutional neural network, a long short-term memory network, and a combined CNN and LSTM network for classifying the four motor imagery direction tasks. The CNN–LSTM achieved the highest mean classification accuracy (87.14%) across four directional classes (left, right, front, back), outperforming the CNN (83.79%) and LSTM (79.74%), with the best Information Transfer Rate of 57.5 bits/min. Results confirm that joint spatial–temporal modeling significantly enhances MI decoding under low-channel constraints, demonstrating the viability of consumer-grade EEG for scalable, real-world BCI applications.

1. Introduction

Brain Computer Interface (BCI) systems using electroencephalography (EEG) have continued to generate intense research interest owing to the potential offered by these systems for the direct interfacing of the human brain with other machines and devices without the need for muscular activity [1]-[5]. Within the different paradigms developed using EEG systems, the Motor Imagery (MI) process, has become a central paradigm for the control of BCIs in neurorehabilitation systems [6], [7], assistive systems like wheelchairs, robots, and other assistive technologies [8], [9], and human-machine interaction systems [10].

One of the major challenges for MI related EEG systems is the balance between signal quality and real-world feasibility [11]. In conventional high-performance MI research, EEG systems with a high density of electrodes, complicated preparation, and lab settings are usually employed [12], [13]. Even though these arrangements are able to record rich spatial information with a high accuracy rate, due to time-containing, user comfort, and high operational costs, these setups become less practical in real world scenarios [14], [15]. High-density clinical systems have greater spatial resolution on the sensorimotor areas (e.g., C3/C4). This allows discriminative patterns to emerge within a short time window. On the other hand, low-channel consumer-grade equipment like the Emotiv Insight has reduced spatial coverage and increased levels of noise. The limitations of the MI-based BCI system have to be addressed with the need for consumer-grade EEG equipment and effective processing pipelines that work well with low-channel equipment.

Research involving low-channel consumer-grade EEG devices has demonstrated the feasibility of MI recognition under constrained settings, but several gaps remain. The EEG data from low-channel systems is more vulnerable to physiological artifacts, drift, and environmental noise, which can make task-related patterns MASKED [16]. Reliable motor imagery (MI) classification depends on effective EEG preprocessing and feature representation, particularly when signals are acquired using consumer-grade devices [17], [18] . Despite emerging studies using low-channel EEG devices, several research gaps remain. First, many studies involve a limited number of participants, restricting evaluation of inter-subject variability. Second, preprocessing and temporal windowing strategies for low-channel EEG system are not systematically analyzed. Third, the effect of temporal window length on multi-class MI performance under constrained spatial resolution has not been comprehensively evaluated in multi-subject consumer-grade datasets.

The major aim of the proposed work is to investigate multi-directional motor imagery classification using a five-channel consumer-grade EEG system. The work provides a systematic analysis of temporal window length and feature dimensionality effects in a multi-subject setting. The major limitation arises due to the fact that low-channel systems tend to have low spatial resolutions. The research contributions of this study are summarized as follows.

- Firstly, collection and analysis of a multi-class motor imagery dataset from 53 participants using a consumer-grade five-channel Emotiv headset
- Secondly, systematic evaluation of temporal window length effects on classification performance under low spatial resolution constraints
- Comparative assessment of CNN, LSTM, and hybrid CNN–LSTM architectures for multi-class MI decoding.
- Quantitative analysis of spatial–temporal feature integration for low-channel EEG systems.

Overall, this work clarifies the role of temporal integration and model design in enhancing classification robustness under low-density EEG conditions and presents an end-to-end framework for motor imagery recognition using consumer-grade hardware and deep learning approaches.

2. Related Work

The current technological drift has focused on the development of lower-cost, non-invasive, and low-intrusive wearable headbands for BCI development [19]. Unlike conventional systems that demand bulky caps, gel conductors, and extensive wired sensors, consumer-grade headbands provide advantages of dry or semi-dry electrodes and wireless connectivity, such as Bluetooth, which make these headbands amenable for ecologically valid experiments for gaming, health, and neuro-rehabilitation applications [20]. Though these headbands tend to offer lower electrode counts, mostly appearing on the central-frontal and temporal areas, these systems have been proved for Motor Imagery (MI) signal processing for controlling systems based on MI of limbs. Several consumer-grade EEG systems have been increasingly used for MI experiments as cited in Table 1.

Muse 2 Headband, popular choice for low-intrusive monitoring, featuring 4 electrodes (TP9, AF7, AF8, TP10). It has been applied so far to reduce response times in the detection of MI with the use of Deep Learning architectures, such as CNN-LSTM, leading to more than 83.8% accuracy for response windows of 2 seconds [21], [22]. Emotiv Series (EPOC X, Insight): EPOC X offers 14 channels, thus providing greater spatial resolution compared to the Muse device [21]. On the other hand, it lacks central electrodes, such as C3/C4, normally used in MI research, which may limit its ability to precisely capture sensorimotor rhythms [22]. The Emotiv Insight (5 channels) has been utilized for real-time cursor control, demonstrating the feasibility of using low-density arrays for continuous navigation tasks. OpenBCI (Cyton + Daisy): It can be identified based on its open-source design, which can record up to 16 channels. It is often used in more complex research, such as the assessment of olfactory perceptual ability and the development of individualized dynamic networks in hand grasping MI [23].

Table 1. Specifications of popular consumer-grade EEG devices [24]

Manufacturer	Model	Release Year	Channels	Sample rate (Hz)
Emotiv	EPOC+	2013	14	128 or 256
MyndPlay	Mindband	2014	1	512
OpenBCI	Ganglion board	2014	4	200
OpenBCI	Cyton board	2014	8	250
OpenBCI	Cyton-Daisy board	2014	16	125
Emotiv	Insight	2015	5	128
InteraXon	Muse 2	2016	4	256
NeuroSky	MindWave Mobile 2	2018	1	512
Emotiv	EPOC-X	2019	14	128 or 256

Motor imagery (MI)–based brain–computer interface (BCI) research has been largely driven by studies using research-grade, high-density EEG systems and publicly available benchmark datasets [25]-[27]. Well-known datasets, such as those from the BCI Competition series and PhysioNet [28], [29], primarily rely on research-grade EEG systems with high channel density and controlled laboratory conditions. While these resources have significantly advanced algorithm development [30]-[32], they do not fully represent the constraints of consumer-grade EEG devices, which are increasingly adopted due to their affordability, portability, and suitability for real-world and rehabilitation-oriented applications [33]. EEG signals acquired from consumer-grade devices are inherently affected by reduced channel counts and lower signal-to-noise ratios [34], making them more susceptible to physiological and environmental artifacts.

This section surveys the existing studies that explore the classification of MI using consumer-grade or low-channel EEG devices, with a focus on three important aspects: (i) the feasibility of MI decoding under the restricted sensing modality, (ii) the preprocessing and feature representation methods to address the issue of noise and restricted spatial information, and (iii) deep learning modeling of MI using consumer-grade EEG devices.

2.1. Feasibility of MI Decoding Under Constrained Sensing Conditions

Traditional BCI research has utilized clinical-grade devices that feature high electrode density (e.g., 32 to 256 channels) combined with conductive gels in order to assure the integrity of signals detected by such electrodes [35]. By contrast, consumer-grade headbands often put a higher value on usability, portability, and low cost, using a small number (between 1 and 14, generally) of dry or semi-dry electrodes [36]. It is despite these hardware limitations that MI can be decoded because motor intentions may be recorded through sensors that are placed on both frontal and temporal lobes of the brain, which are involved in controlling movement and sensory processing [37].

While conventional BCI studies utilize clinical-grade equipment with a high number of electrodes (32–256 channels) and gel electrodes for signal quality, consumer-grade headbands emphasize ease of use, portability, and cost-effectiveness. These headbands use a small number of dry electrodes ranging from 1 to 14 electrodes. Although the equipment is not adequate for BCI signal detection, it is still possible to decode MI signals as the signals from the sensors placed on the frontal and temporal lobes can detect motor intentions. Research using the Muse 2 Headband concluded that binary MI signal classification tasks, such as distinguishing the use of the right hand from the left hand, can be achieved at an accuracy of over 83% without stimulating the primary motor cortex at C3/C4. However, the major challenge still exists in BC Illiteracy, where about 15%-30% of the population is not capable of generating differential signals independently of their environment for signal detection purposes [38]. A concept of Embodiment-based Virtual Reality (VR) using avatars for improving MI signal participation and speed was introduced in [39].

2.2. Preprocessing and Feature Representation Strategies

Consumer EEG applications require mitigation of noise in non-shielded ecological environments [35]. Standard preprocessing involves Notch filtering (50/60 Hz) to remove line interference and band-pass filtering (commonly 8-30 Hz) that isolates mu and beta rhythms relevant to the motor tasks [40], [41]. FastICACorr, ICA, and ASR have been increasingly applied to peel away ocular and muscle artifacts without compromising the underlying neural signal. For feature representation, CSP remains a benchmark but is often enhanced for low-channel arrays by Filter Banks or wavelet-based approaches, which capture intra-band temporal and frequency discriminators. Recent research has also introduced Morphological Pattern Spectrum (MPS) for the encoding of potential variations to provide a more comprehensive view of dynamic neural networks that are established during MI. Studies as recent as 2024 have employed the 14-channel Emotiv Epoc X for the analysis towards rehabilitation, demonstrating the differentiation between motor execution and motor imagery for six different actions by extracting band power features from the theta to gamma bands [42]. In addition, benchmarking experiments involving the Emotiv Epoc X have covered the binary differentiation between MI and Motor Execution (ME) by 1D features with baseline normalization for setting up baselines on classical machine learning performance [43].

[44] showed that CSP and LDA can be used for MI-based control of a lower-limb exoskeleton, while [45] confirmed the clinical validity of MI–BCI rehabilitation in stroke patients. Within the context of EEG feature extraction, [46] suggested the use of a hybrid ICA–WT–CSP pipeline in order to enhance artifact removal and spatial discrimination. Although this method attained higher accuracy compared to classical methods, it relies on handcrafted feature engineering and several sequential processing steps. [47] utilized a GA in order to execute subject-specific feature selection in hybrid BCI systems and demonstrated great performance on multi-modal datasets. However, GA-driven optimization does add extra computational cost and is inherently subject-dependent.

2.3. Deep Learning–Based Modeling Approaches

This paradigm has supported end-to-end classification, where unprocessed data is directly inputted into a neural network for feature extraction and intention classification concurrently [19], [48]. The Convolutional Neural Network (CNNs) architecture has been the leading approach, thanks to their potential for spatial-spectral feature extraction from a small set of electrode recordings [49]. To process the non-stationary, time-series nature of the EEG, Recurrent Neural Networks (RNNs), as well as Long Short-Term Memory (LSTM) architectures, have been used for temporal pattern extraction. More importantly, Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs) have been employed for data augmentation for addressing the "small-sample learning" problem of training individualized BCIs. Moreover, hybrid architectures such as "CNN-LSTM" have successfully shortened BCI reaction times to under 2 seconds with accuracy above 80%, which is an essential need for real-time processing of engineering applications. Other resilient MI-associated decoding schemes have utilized the 15-channel OpenBCI platform for MI/rest classification from unfiltered time-series data, where EMG-based rejection has been used for ensuring data cleanliness [50].

[51] developed a domain adaptation modeling method based on transformers that equalizes feature distributions between subjects, achieving excellent generalization capabilities on public datasets. Nevertheless, this technique demands extensive training datasets, heavy computational power, and complex adversarial training, which makes it inappropriate for resource-constrained or real-time BCI applications. Similarly, [52] developed EDANet by integrating domain-aware spatial

filters with attention mechanisms for temporal modeling. From an application side, MI-BCIs have been effectively applied to robotic control tasks and also to rehabilitation applications [53]. To be specific, the potential of using MI control for controlling a lower-limb exoskeleton using CSP and LDA was shown in [54]. The efficacy of using MI-BCI-based rehabilitation in re-habilitating stroke patients was, in turn, established in [45]. Though these papers demonstrate that there is potential for using MI-BCIs, which have been confirmed to be practically useful, classification accuracy in these studies remains relatively low, and they typically employ traditional signal processing methods that are highly sensitive to noise and inter_subject variability. In addition, such systems also demand extensive training for users, with the instability of the EEG signals not being taken into account. Recently, imaging paradigms have been proposed as methods to increase the discriminability of EEG signals. In 2009, motor imaging and speech imaging were combined to gain higher accuracy of classification as well as a stronger event-related desynchronization pattern [55]. While an improvement over other methods, this approach also includes additional cognitive load, which can be challenging for neurological patients [56].

Based on such a foundation, the contributions of our proposed dataset and approach lie within the further advancement of the state of the art for the following reasons: multi-class MI classification task of Left, Right, Front, Back, supported by a highly constrained 5-channel consumer-grade EEG system.

3. Methodology

This section covers the details of the experimental design, data acquisition protocol, EEG preprocessing pipeline, event annotation and epoch extraction strategy, dataset construction, and the approach for deep learning-based modeling. As illustrated in Fig. 1, the end-to-end EEG processing pipeline used in the study, from data acquisition to epoch extraction for motor imagery classification.

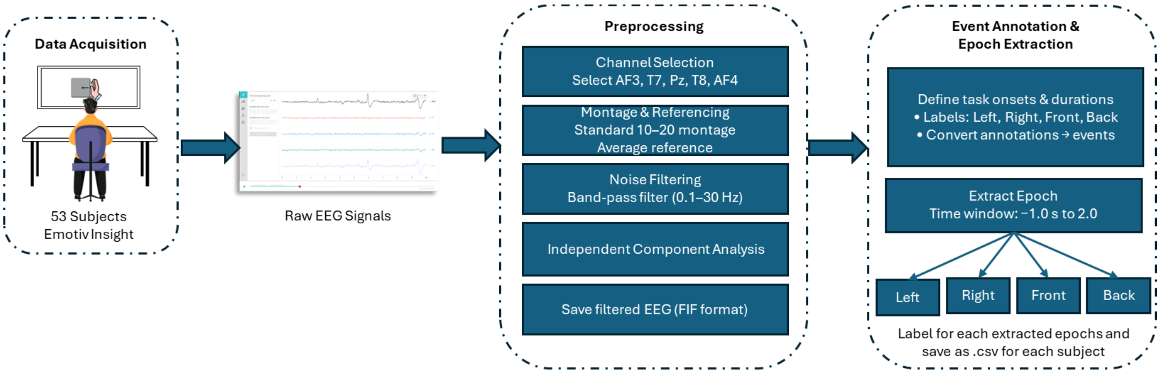


Fig. 1. EEG data acquisition and preprocessing workflow

The EEG signals are first recorded from participants while they are performing video-guided motor imagery tasks using the Emotiv Insight headset device. Then, the raw EEG signals are imported and preprocessed. This involves the application of the 10-20 montage placement with average referencing, band-pass filtering to eliminate slow drifts and high-frequency noise in the signals, independent component analysis to investigate component topography and temporal characteristics to identify gross artifacts before epoch extraction, and the EEG signals are recorded in FIF format to allow reproducibility. Additionally, the task marker is converted to an event marker, and motor imagery classes such as left, right, front, and back are labeled and segmented into epochs with the objective of structuring the input to deep learning classification methods.

Following the extracted epochs, Fig. 2 shows the subsequent feature aggregation and sequence reconstruction and modeling process. The extracted EEG epochs, sampled at 128 Hz and selected from five channels, undergo temporal aggregation using mean pooling to reduce noise and increase the stability of the signal representation. The aggregated signals are then converted to fixed-length sliding windows to form multivariate temporal sequences, which maintain the temporal dependencies in the data. The multivariate temporal sequences are split into training, validation, and test sets and used to feed deep learning approaches such as CNN, LSTM, and hybrid CNN-LSTM to systematically evaluate the effectiveness of spatial, temporal, and spatio-temporal learning approaches for multi-directional motor imagery classification.

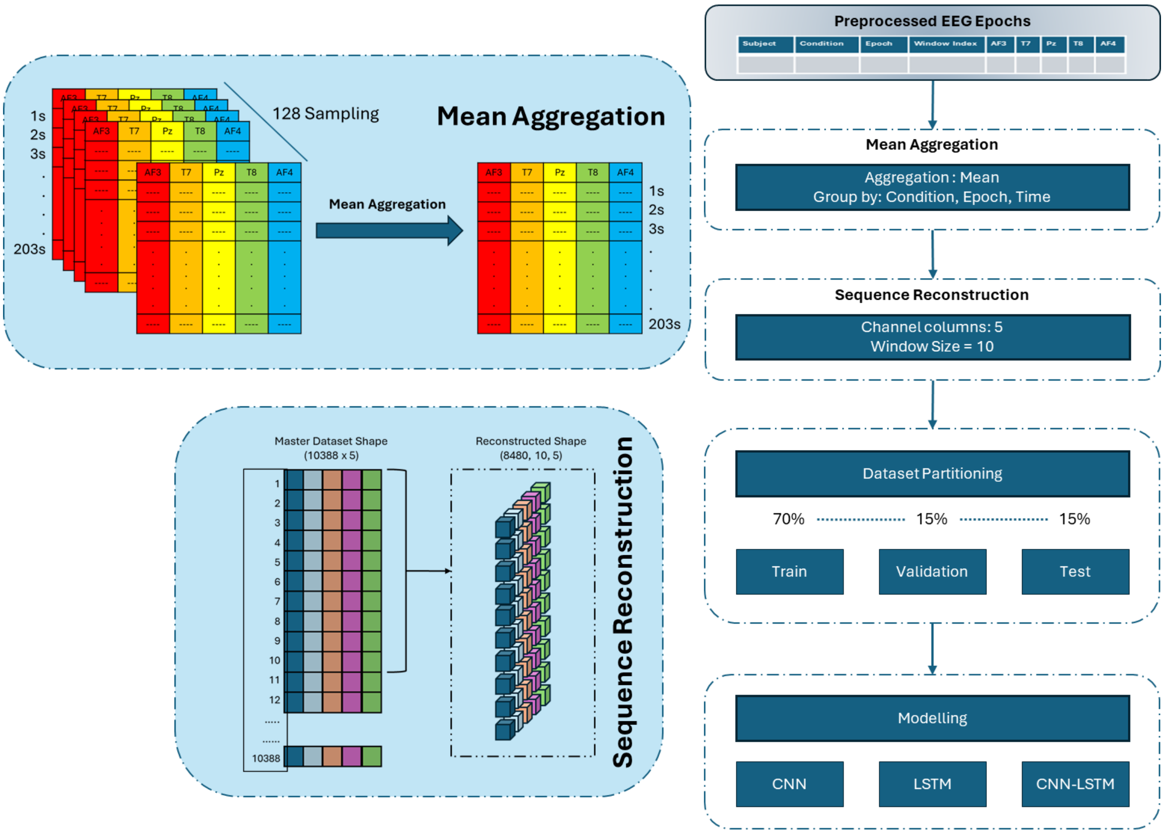


Fig. 2. Feature aggregation, sequence reconstruction, and deep learning modeling pipeline

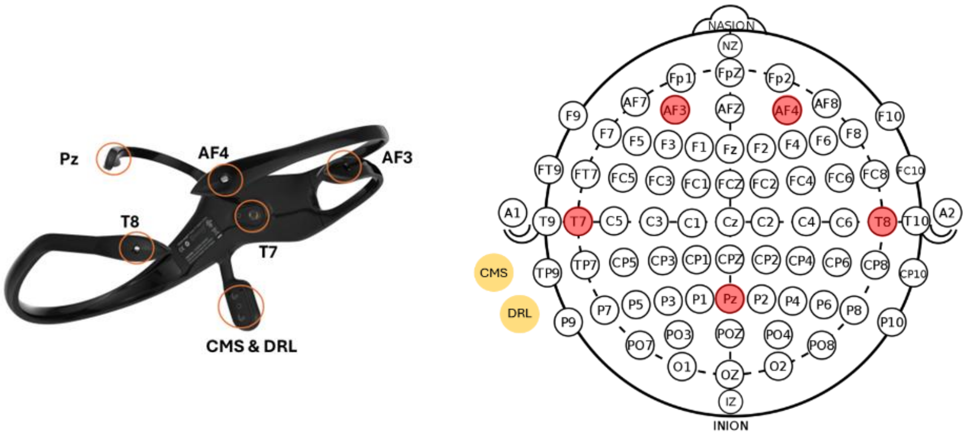
3.1. Data Acquisition

3.1.1. Participants and Experimental Setup

EEG data was collected from a total of 53 participants using the Emotiv Insight device, as the participants sat comfortably in front of the display monitor in a quiet indoor environment. The headset was adjusted to ensure stable contact of the electrodes before the experiment, as recommended by the manufacturer. Fig. 3 shows the hardware configuration of the EEG equipment used in the experiment. Fig. 3 (a) shows the Emotiv Insight consumer-grade EEG headset, which consists of five active electrodes (AF3, AF4, T7, T8, and Pz) as well as the reference electrodes Common Mode Sense (CMS) and Driven Right Leg (DRL). Fig. 3 (b) shows the placement of the active and reference electrodes of the equipment on the head, as defined by the international 10-20 system, showing the anatomical correspondence of the active electrodes to the frontal, temporal, and parietal areas of the scalp, which are associated with motor imagery activity, even though the equipment has only five active electrodes.

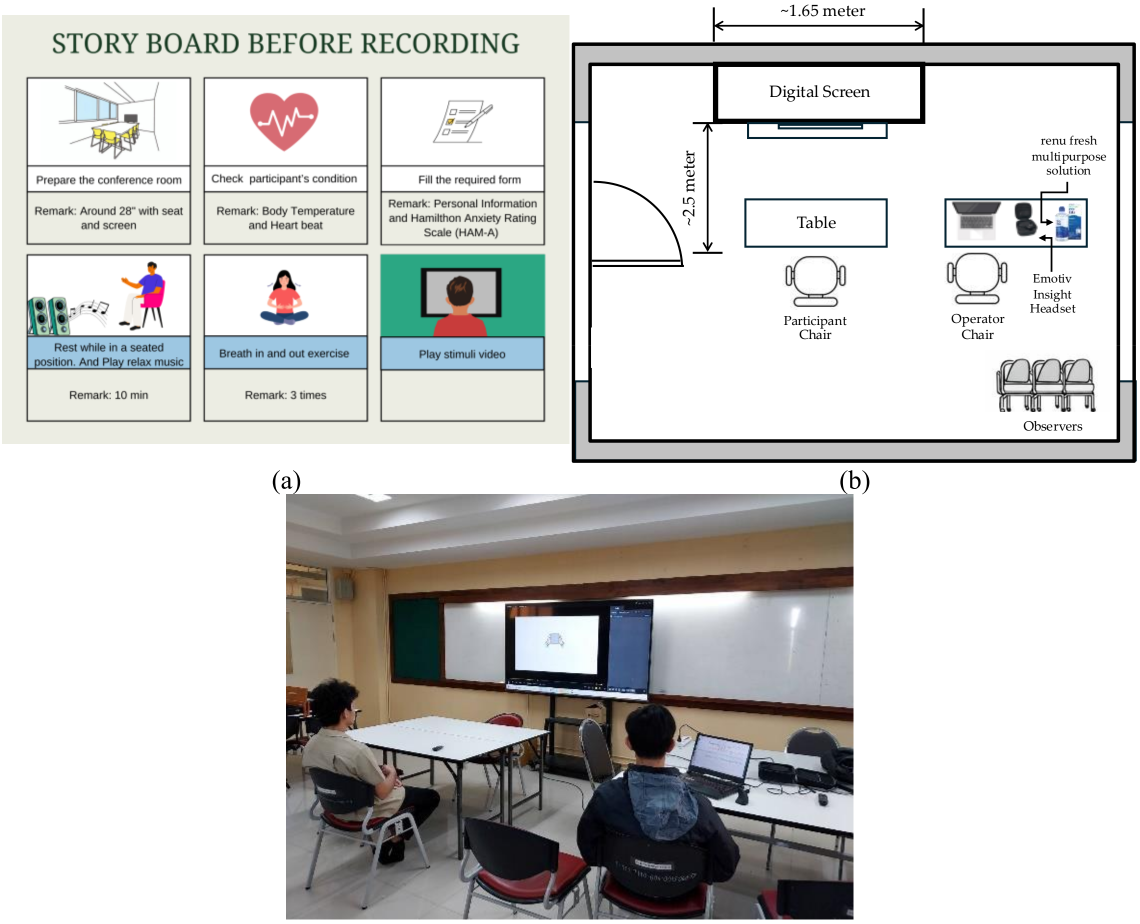
3.1.2. Motor Imagery Task Design

The participants were asked to do motor imagery tasks in reaction to pre-recorded videos of stimuli specially designed to elicit imagery of one of the four directional categories, namely LEFT, RIGHT, FRONT, or BACK. The visual stimuli provided unambiguous task cues with precise temporal correspondence between stimulus delivery and the neural responses, as direct responses to stimuli were recorded using EEG. No muscle movements were allowed during the EEG recording process. Markers were recorded along with the EEG signal indicating the delivery of stimuli and completion of the task time for the event markers. These markers were then used to label the events for the supervised phase of the process. Fig. 4 outlines the design of the motor imagery task process followed during the EEG signal Acquisition process.



(a) (b)

Fig. 3. Emotiv Insight EEG headset and electrode configuration: (a) Physical layout of the Emotiv Insight headset showing the five recording electrodes along with the reference electrodes (CMS and DRL); (b) Corresponding electrode positions highlighted on the international 10–20 EEG system



(a) (b)

(c)

Fig. 4. Motor imagery task design and experimental setup: (a) Storyboard of the pre-recording protocol; (b) Experimental setup/ layout design; (c) Experimental setup during EEG recording

As can be observed from Fig. 4 (a), a storyboard was employed to standardize the participants' preparation before the recordings, including room preparation, participants' conditions check, informed consent, relaxation techniques, breathing exercises, and familiarization of the participants to the video materials. Participants' preparation was further employed to standardize the experimental conditions for better preparedness of the participants before every session. The experimental layout design is presented in Fig. 4 (b), which can be considered as a guideline for data acquisition during the experiment to standardize the placement of the headset on the participant's head and the experimental conditions employed during data acquisition. The real experimental environment can be observed from Fig. 4 (c), which was designed to record the participants performing motor imagery tasks based on video instructions by being seated comfortably on the chair equipped with an EEG headset to standardize the experimental conditions employed to record the data from the participants.

Table 2 reports participant demographics, duration of recording, configuration of the task, EEG hardware characteristics, and acquisition settings. Jointly these parameters determine the scale and the structure of the dataset. As depicted in Fig. 5, the resulting dataset has a good balance among all four motor imagery classes: left, right, front, and back. This balanced distribution among the classes reduces sampling bias and ensures that the performance of the CNN, LSTM, and CNN–LSTM models indicates true discriminative capability among classes during training, validation, and testing.

Table 2. Overall classification performance across models

Parameter	Value
Subjects	53
Age	22-25
Run length	0s -203s
Tasks	4
Electrodes	5
Sampling rate (Hz.)	128

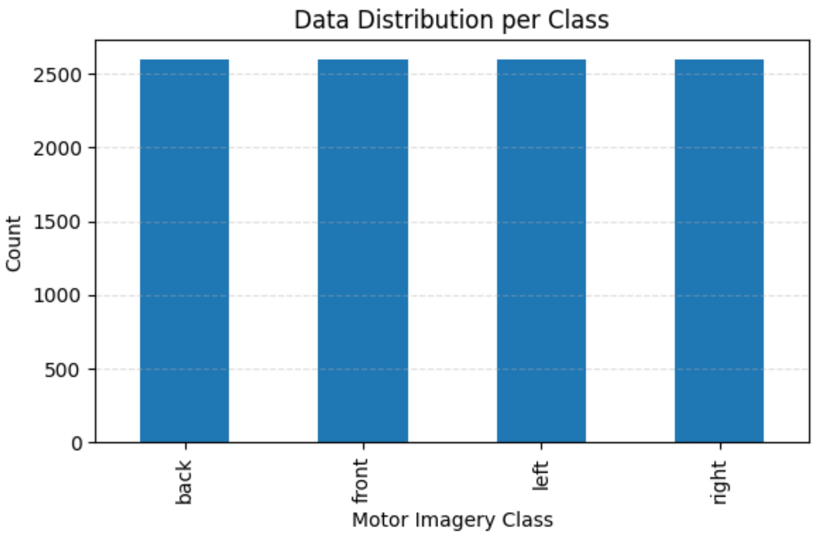


Fig. 5. Distribution of EEG samples across motor imagery classes

3.2. EEG Preprocessing

3.2.1. Channel Selection and Referencing

EEG preprocessing was done using the MNE-Python environment. It provides a standardized, reproducible set of tools for EEG/MEG signal processing. The Emotiv Insight headset includes five EEG channels placed according to the international 10-20 placement system at AF3, T7, Pz, T8, and AF4. These five channels have been selected here, as they cover frontal, temporal, and parietal regions commonly reported to be involved in the generation of motor imagery-related neural activity. The Emotiv Insight headset employs a CMS/DRL active reference system positioned on the left mastoid process, which differs from a traditional passive mastoid reference (A1/A2). To reduce dependence on the hardware reference and improve inter-channel comparability, signals were re-referenced offline to a Common Average Reference (CAR). Although the five-channel configuration limits full spatial averaging, CAR provides a balanced referencing approach for low-density wearable EEG systems. This is a common approach that is widely adopted in EEG experiments with few channels; it also enhances cross-subject comparability.

We formally denote the raw EEG signal recorded at channel i at time t as let $x_i(t)$, where i= 1,2,…,N and N is the total number of EEG channels. Let $\bar{x}(t)$ be the average reference signal computed as the arithmetic mean of all channel signals at each time instant as given in Equation (1):

$$\bar{x}(t) = \frac{1}{N}\sum_{i=1}^{N} x_i\,(t) \tag{1}$$

For each communication channel, it is then re-referenced based on the subtraction of this mean signal as shown in Equation.

$$x_i^{ref}(t) = x_i(t) - \bar{x}(t) \tag{2}$$

where, $x_i^{\mathrm{ref}}(t)$ is referred to as the re-referenced EEG signal on channel i, $x_i(t)$ is referred to as EEG signal on channel i, and $\bar{x}(t)$ is referred.

3.2.2. Filtering and Noise Reduction

To eliminate low-frequency drift and high-frequency noise, the EEG signals were band-pass filtered between 0.1 Hz and 30 Hz using a Finite Impulse Response (FIR) filter. This frequency band is effective in maintaining rhythm-related frequencies in motor imagery that involve mu and beta frequencies, eliminating low frequencies attributed to artifacts and high-frequency noise.

Independent Component Analysis (ICA) was applied as an exploratory signal inspection step. Because the number of independent components is mathematically limited by the number of recording channels, the five-channel configuration yielded five ICA components. ICA was not used for automated artifact rejection in this study.

Component scalp topographies and temporal dynamics were visually analyzed to identify potential gross artifacts prior to epoch extraction on the basis of spatial distributions and time-course as per predefined rules. All components were retained to avoid inadvertent removal of task-relevant neural information. Filtered EEG data was stored in the FIF file format, maintaining all data features and markers.

3.3. Event Annotation and Epoch Extraction

3.3.1. Event Definition

Fig. 6 shows the event annotation structure for the extraction of EEG epochs. The task order is divided into carefully distinguished phases which consist of a pre-rest period (PRE), motor imagery or task execution (10 s), and a post-rest period or task completion (COM). The baseline period consists of the beginning of the task-related recording phases which alternate between the eyes-closed and eyes-open states, then consist of repeated trials for the directional motor imagery (left, right, front, and back). The motor imagery classes are replicated multiple times with fixed time durations, allowing for the exact transformation of the task annotations to event markers for the extraction of epochs. The task-related annotations obtained during the data acquisition were transformed to discrete event markers. Each event was marked to identify the beginning of the directional motor imagery task. The events are divided into four classes, namely left, right, front, and back.

3.3.2. Epoching Strategy

EEG epochs were extracted from −1.0 s to 2.0 s relative to stimulus onset, with baseline correction applied using the −0.5 to 0 s interval. Epochs exceeding ±200 µV were rejected. The prestimulation period was used as the baseline period; the post-stimulation period was used to derive the task-related activities of the brain linked with the execution of motor imagery. After the epoch extraction process, epochs that were dominated by noise or artifacts were excluded using pre-defined thresholds on the epoch amplitudes and variances to ensure quality in the data post-processing.

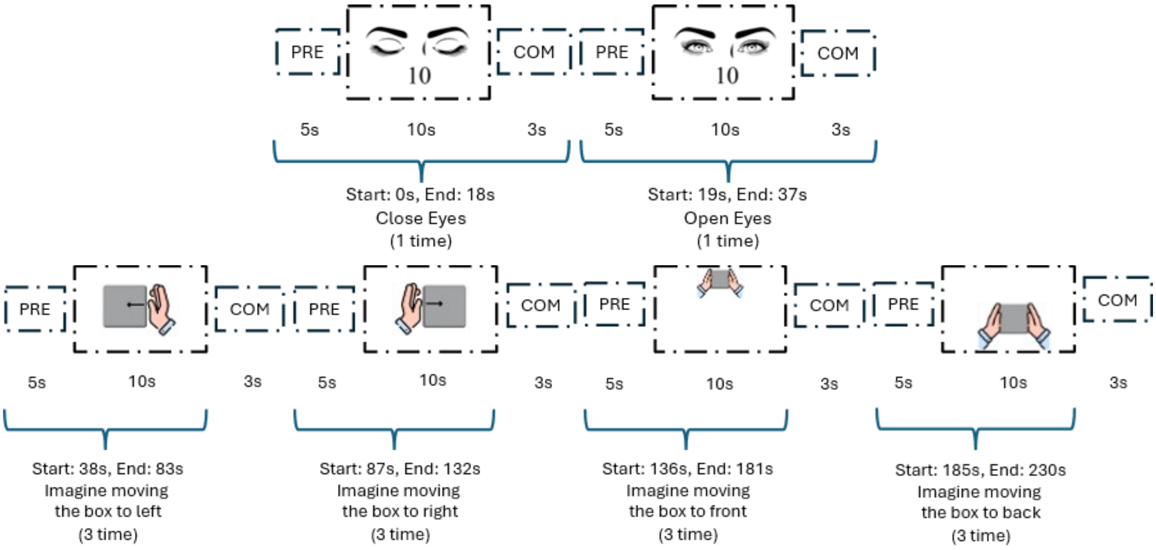


Fig. 6. Event definition and temporal structure for EEG epoch extraction

To describe the epoched EEG data, let $X(t) \in \mathbb{R}^N$ denote the multichannel EEG signal recorded at time t, where N is the number of EEG channels. After sampling at frequency f_s, the discrete EEG signal is (Equation (3)):

$$X[n] = X\left(\frac{n}{f_s}\right) \tag{3}$$

The experiment consists of predefined task intervals (Equation (4)):

$$\mathcal{T} = \{(t_s^{(k)}, t_e^{(k)}, c^{(k)})\}_{k=1}^{K} \tag{4}$$

where $t_s^{(k)}$ and $t_e^{(k)}$ denote the start and end times of task k, and $c^{(k)}$ is the corresponding task label. In this study, tasks correspond to eye-state and motor imagery conditions.

For each task k, EEG samples are extracted as an epoch as shown in Equation (5):

$$E^{(k)} = \left\{X(t) \mid t \in [t_s^{(k)}, t_e^{(k)}]\right\} \tag{5}$$

In discrete time, this becomes (Equation (6)):

$$E^{(k)} \in \mathbb{R}^{N \times L_k}, L_k = (t_e^{(k)} - t_s^{(k)}) f_s \tag{6}$$

Each epoch $E^{(k)}$ is assigned the label $c^{(k)}$ and used as input for subsequent aggregation and sequence modeling.

3.4. Temporal Aggregation and Dataset Construction

3.4.1. Mean Aggregation Across Time

Each extracted epoch consisted of multi-channel EEG signals sampled at 128 Hz. To reduce high-frequency variability and emphasize task-relevant temporal patterns, mean aggregation was applied over short, non-overlapping time segments. Specifically, signal samples within each temporal segment were averaged to produce a smoothed representation of neural activity across channels. Let $E_k(i,t)$denote the signal of channel $i$at time index $t$within epoch k. The mean-aggregated signal $\tilde{E}_k$is computed as Equation (7):

$$\tilde{E}_k(i,s) = \frac{1}{|\Delta|}\sum_{t\in\Delta_s} E_k\,(i,t), s = 1,2,\dots,S \tag{7}$$

where:

- $\Delta_s$is the set of samples in segment s,
- $\tilde{E}_k \in \mathbb{R}^{N\times S}$.

This aggregation approach resulted in a reduction of the data dimensions with respect to time, and thus produced features amenable for modeling as a sequence.

3.4.2. Windowing and Sequence Reconstruction

After the processes of data preprocessing and extraction of epoch data, EEG features for both aggregated data sets have been formed into fixed-size temporal templates to capture motor imagery activity characteristics in time. Let $\mathbf{x}_t \in \mathbb{R}^C$denote the EEG feature vector at time step t, where $C = 5$corresponds to the selected EEG channels (AF3, T7, Pz, T8, AF4). A temporal window of length $W = 10$time steps was constructed as.

$$\mathrm{X}_i = [\mathrm{x}_i, \mathrm{x}_{i+1}, \dots, \mathrm{x}_{i+W-1}], \mathrm{X}_i \in \mathbb{R}^{W\times C}. \tag{8}$$

This sliding window technique enabled the conversion of the continuous EEG signal to a sequence representation amenable to deep learning algorithms. This helped maintain the sequential nature of the EEG signal samples, as well as a reduced dimensionality representation of the original signal by aggregating the means.

Next, sequential reconstruction was used to produce structured input sequences for recurrent and hybrid models. The resultant dataset was denoted by.

$$\mathcal{D} = \{(\mathrm{X}_i, y_i)\}_{i=1}^{N}, \tag{9}$$

where $y_i \in \{1,2,3,4\}$ is the motor imagery class label for left, right, front, and back motor imagery, respectively, and N is the total number of reconstructed sequences.

3.4.3. Master Dataset Formation

In order to support both intra- and intersubject learning, the EEG sequences of all subjects were pooled together to create a master dataset. Each sequence was identified with a unique identifier for the subject, the type of task, the epoch, and the window identifier for the sequence.

The final master dataset, referred to as $\mathcal{D}_{\text{master}}$, was composed of labeled sequences for each of the four motor imagery classes and for all 53 subjects. The common format facilitated experimentation with both single-subject and multi-subject training paradigms.

3.5. Dataset Partitioning

The master dataset was split into training, validation, and testing sets to have an unbiased assessment of its performance. The data was split on the sequence level and in a way that it is well distributed in each set. It is important to note that this constitutes a subject-dependent evaluation strategy, as samples from the same participant may appear in multiple subsets. This approach was selected to maximize data utilization and to assess model performance under a multi-subject pooled training paradigm.

The dataset was partitioned at the sequence level using a stratified split of 70% for training, 15% for validation, and 15% for testing to preserve class balance across subsets. Apart from subject agnostic, tasks were also performed using the combination strategy for multi-subject training. This is because, presumably, greater data diversity will have a positive effect on the generalization performance.

3.6. Deep Learning Models

Three deep learning architectures were developed and evaluated to decode multi-directional motor imagery from low-channel EEG sequences: a Convolutional Neural Network (CNN), a Long Short-Term Memory (LSTM) network, and a hybrid CNN–LSTM model.

3.6.1. Convolutional Neural Network (CNN)

The CNN architecture is designed to extract spatial temporal representations from the multichannel EEG sequences. One-dimensional convolutional layers are applied along the temporal axis to learn localized patterns within short temporal segments while jointly integrating information across channels. The calculating function for a convolutional layer could be described in:

$$\mathrm{h}^{(l)} = f(\mathrm{W}^{(l)} * \mathrm{h}^{(l-1)} + \mathrm{b}^{(l)}), \tag{10}$$

where $\mathrm{W}^{(l)}$ and $\mathrm{b}^{(l)}$ are learnable weights and biases for the l-th layer, and $f(\cdot)$ is a nonlinear activation function.

Batch normalization is applied after each convolution to stabilize training and accelerate convergence. Max pooling is used to reduce temporal dimensionality and improve feature robustness. A global average pooling layer aggregates high-level feature maps before they are passed to fully connected layers for classification. The final dense layer with softmax activation outputs class probabilities for the four motor imagery directions as illustrated in Fig. 7.

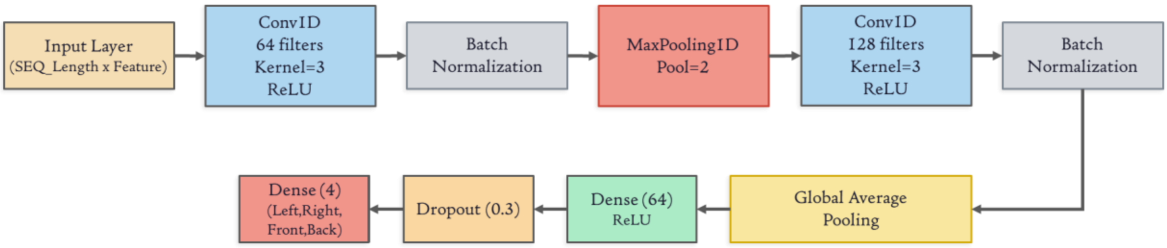


Fig. 7. CNN model architecture

3.6.2. Long Short-Term Memory Network (LSTM)

In their design, the LSTM architecture concentrated on the exploitation of long range dependencies in the EEG signal for motor imagery. This allowed the sequences to be processed in a serial manner where the LSTM architecture leveraged the use of gated memories for retaining information in the sequences. This process for the LSTM architecture at time step t is represented by Equation (11). In the equation, $\mathbf{h}_t$ and $\mathbf{x}_t$ are the hidden units and the input vector, respectively, at time *t*.

$$\mathrm{h}_t = \mathrm{LSTM}(\mathrm{x}_t, \mathrm{h}_{t-1}), \tag{11}$$

The architecture consists of two stacked LSTM layers (128 and 64 units). The first layer returns full temporal sequences, allowing deeper temporal abstraction in the second layer as shown in Fig. 8. Dropout regularization is applied between layers to mitigate overfitting. The resulting representation is mapped to class probabilities through fully connected layers with ReLU activation followed by a softmax output layer.

3.6.3. Hybrid CNN–LSTM Model

The hybrid model based on the CNN and LSTM architectures provides a unified model for spatial feature extraction and temporal modeling. First, the convolutional layer is used to extract discriminative spatial–temporal features from short segments of the EEG data. Then, the extracted high-level feature representations are processed by the LSTM layer to model sequential dependencies in the temporal dimension. The whole process can be described as follows:

$$\mathrm{h}_t = \mathrm{LSTM}(\mathrm{CNN}(\mathrm{X}_t)). \tag{12}$$

By utilizing convolutional filters and recurrent memory cells as shown in Fig. 9, it can be seen that the hybrid makes use of local spatial and temporal contextual cues. This is particularly useful in cases of low-channel EEG configurations, where spatial resolution is not sufficient and temporal integration becomes a crucial factor in improving motor imagery classification.

Fig. 8. LSTM model architecture

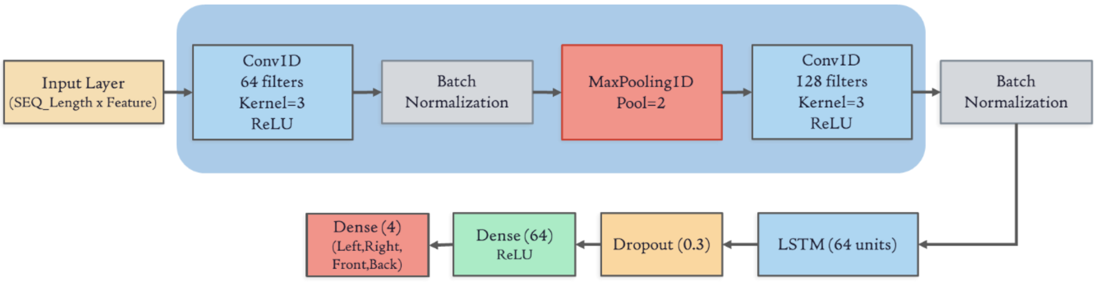


Fig. 9. CNN-LSTM model architecture

3.7. Model Training and Evaluation

All models were trained using supervised learning with categorical cross-entropy loss, defined as.

$$\mathcal{L} = -\sum_{k=1}^{K} y_k \log(\hat{y}_k), \tag{13}$$

where $K = 4$ is the number of motor imagery classes, y_k is the ground-truth label, and $\hat{y}_k$ is the predicted class probability. Optimization algorithms were adaptive gradient-based, with early stopping introduced to eliminate overfitting. The dataset was partitioned into training, validation, and test sets, and all evaluation metrics were computed on the held-out test set.

Model performance was assessed using multiple complementary metrics. Classification accuracy and class-wise accuracy were used to evaluate overall and per-class decoding performance. Information Transfer Rate (ITR) was computed to quantify the effective communication throughput, accounting for both accuracy and decision speed. Total system latency, comprising temporal window length and model inference time, was measured to assess real-time feasibility. Additionally, paired t-tests were conducted between consecutive window sizes for each model to determine whether performance improvements across temporal window lengths were statistically significant ($p < 0.05$).

4. Results

This section includes the results obtained in the experiment using multi-directional motor imagery classification based on low-channel EEG signal processing using the Emotiv Insight EEG headset. The performance of the deep learning architectures CNN, LSTM, and the combined model CNN-LSTM was evaluated on a large multi-subject EEG dataset. The results were obtained at the subject, class, and model levels to provide a comprehensive analysis of the classification behavior and the generalization capability. The performance was evaluated using accuracy, information transfer rate, system latency, and paired t-tests on consecutive window sizes. The results were obtained directly from the implementation without any exclusion or post-processing on the data.

4.1. Independent Component Analysis (ICA) Results and Signal Quality Assessment

Independent Component Analysis (ICA) was applied to the preprocessed EEG data to evaluate signal quality and identify potential artifact-related components prior to epoch modeling. Fig. 10 presents the scalp topography of the first four independent components extracted, namely ICA000, ICA001, ICA002, and ICA003.

ICA000 displays a widespread topography across the frontal-central areas without any focal peak, indicating a global cortical activity without the presence of ocular or muscular artifacts. ICA001 presents a lateralized topography in the right temporal area, indicating a differential cortical activity in the motor imagery paradigm. ICA002 presents a posterior topography that covers the parietal and occipital areas, indicating a possible engagement in visuospatial activities as a consequence of the video paradigm. Lastly, the topography in ICA003 presents frontal activity; however, the smooth distribution without any focal peak makes it less likely to have strong ocular artifacts. Overall, the independent components extracted did not show clear topography related to severe artifacts such as eye blinks and muscle activities. Thus, all the independent components extracted were included in the modeling process.

4.2. Temporal Segment Activity and ERP/ERF Patterns

Fig. 10 depicts the spatial, temporal, and spectral characteristics of the four independent components obtained from the EEG signal using Independent Component Analysis. The figures (a, b, c, d) represent the four independent components obtained from the EEG signal, namely ICA002, ICA001, ICA000, and ICA003. These scalp topographies also show the unique spatial characteristics of each independent component. In ICA002 (a), it is evident that there is a posterior-inferior dominance, which is characteristic of parietal or occipital origin. In ICA001 (b), it is evident that there is a lateralized pattern in the right frontal-temporal region, which may be characteristic of eye movement or muscle activity. In ICA000 (c), it is evident that there is a widespread anterior-posterior gradient with high voltage in the anterior regions, which is characteristic of ocular artifacts. In ICA003 (d), it is evident that there is a unique pattern of activation in the frontal regions, which may be characteristic of eye blink or facial muscle activity shown in Fig. 11.

This can be easily seen in the image and the ERP/ERF curves plotted for each segment. For the ICA002 case, there is a moderate and consistent fluctuation on a lower scale, confirming its neural nature. Furthermore, in the case of the ICA000 and the ICA003, there is a high fluctuation in the signal over time, confirming that it is non-neural in nature. This can also be seen more clearly when the power spectra are used to analyze the results, where a high power at lower frequencies (<4 Hz) is a strong indication of ocular artifacts, and a smooth 1/f curve is a strong indication of neural activity. The variance distributions and the lack of removed parts in the segments (0.00%) show that there is stability in the de-composition as well as the good quality of the data. From all the results obtained in the above analyses, it is possible to make the appropriate judgments on the rejection/retention of the components before the EEG reconstruction.

4.3. Event-Related Topographic Responses

In Fig. 12, the evoked topographies for left, right, front, and back motor imagery tasks around 0.75–0.84 s after stimulus onset are shown. In all these topographies, although the spatial resolution is restricted by the five-channel, consumer-grade EEG setup, a unique distribution for each trial type is clearly discernible. In the left and right motor imagery topographies (Fig. 12 (a) & (b)), the hemispherical asymmetries highlighted by the changes in polarity across the temporal electrodes follow established patterns for contralateral motor imagery representations. In contrast, the anterior-posterior distributed patterns in the topologies for the front and back imagery tasks (Fig. 12 (c) & (d)) reflect higher involvement of frontal-parietal networks with less laterality, suggesting that the relatively poorer accuracy for the former imagery type in the experimental results might be due to relatively diffuse activations. However, these topologies must be viewed as indicative of relative distributions among the electrodes rather than actual localization in the cortex. Nonetheless, the presence of class-specific and reproducible topographic differences confirms that meaningful directional motor imagery information is preserved in low-channel EEG recordings and provides a neurophysiological basis for the observed deep learning classification performance.

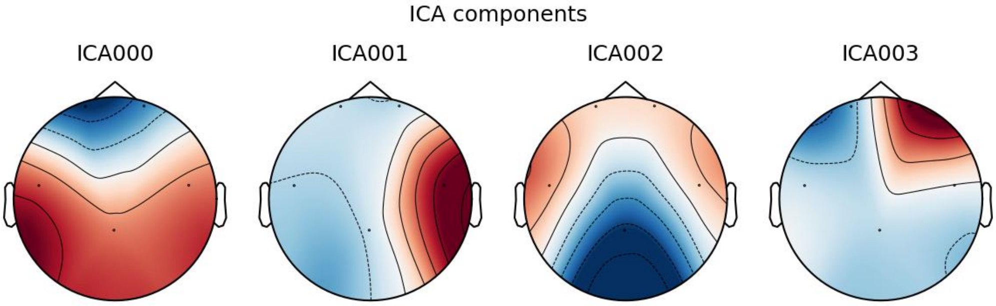
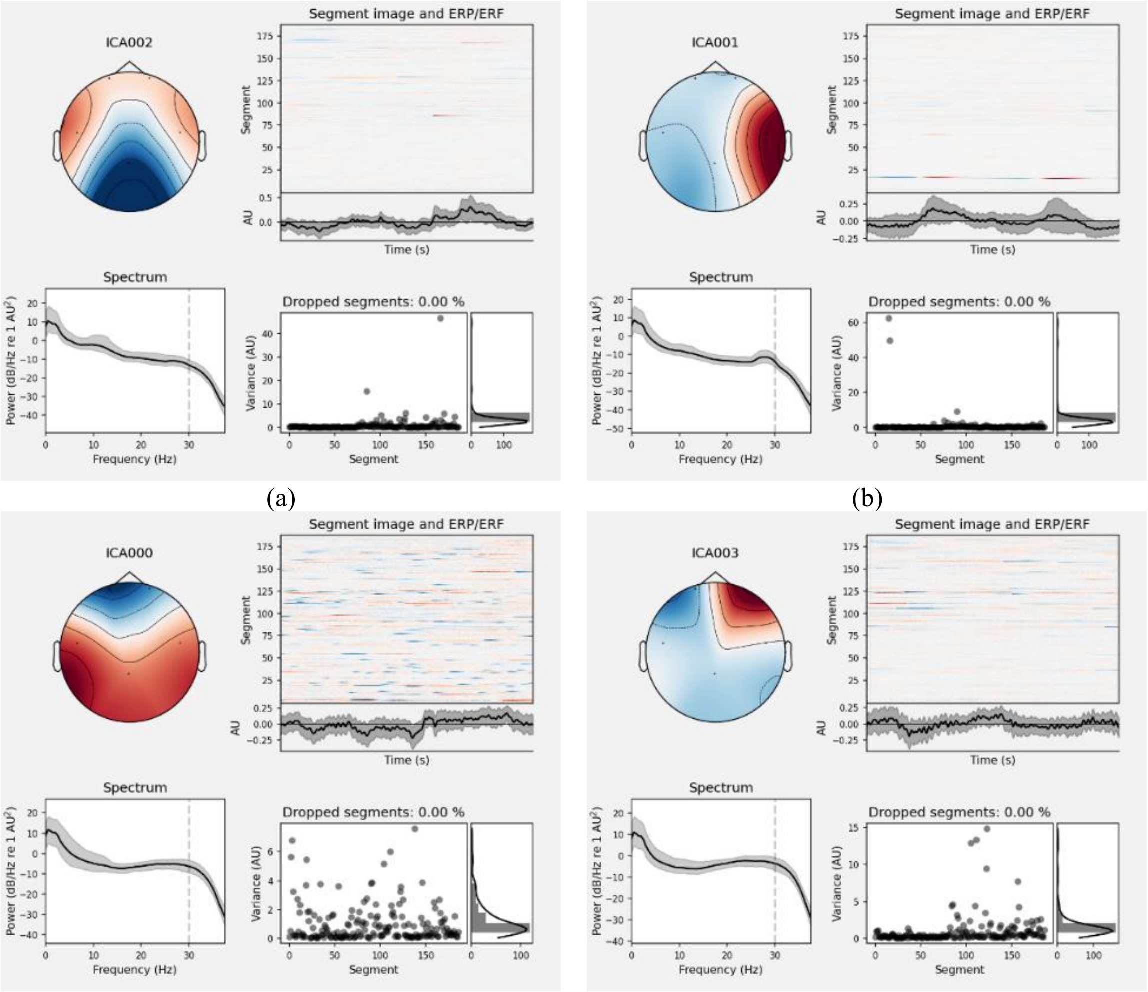


Fig. 10. ICA components



(a) (b)

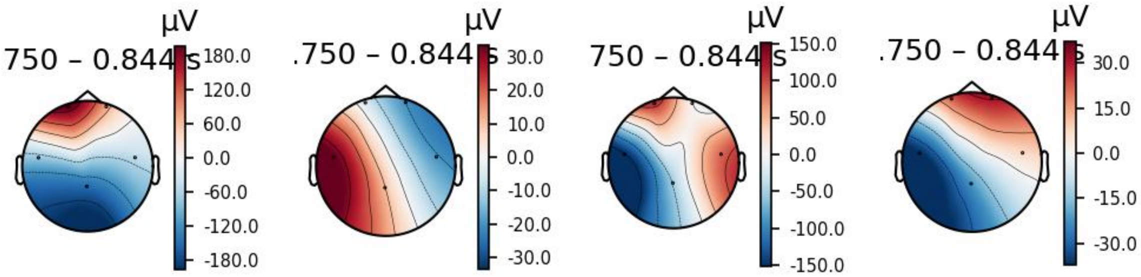
(c) (d)

Fig. 11. Independent component analysis (ICA) Decomposition and component evaluation: (a) ICA002; (b) ICA001; (c) ICA000; (d) ICA003

4.4. Overview of Experimental Outcomes

A total of 53 subjects were included in the analysis, resulting in 159 trained models (53 subjects × 3 architectures). For each subject and model, classification performance was evaluated across four motor imagery classes: left, right, front, and back. Class-wise accuracies were computed and further aggregated to obtain a mean accuracy score per subject–model combination.

Fig. 13 presents the confusion matrices for the (a) LSTM, (b) CNN, and (c) hybrid CNN–LSTM models across the four motor imagery classes (back, front, left, and right). In Fig. 13 (a), the LSTM model shows moderate diagonal dominance with noticeable confusion between neighboring directional classes, indicating that temporal modeling alone is insufficient to fully separate similar motor imagery patterns. In Fig. 13 (b), the CNN model improves class discrimination, particularly for the front and right classes, although some misclassification remains between left and right directions, likely due to limited spatial resolution in the low-channel EEG setup. As shown in Fig. 13 (c), the hybrid CNN–LSTM model exhibits the strongest diagonal dominance and the lowest off-diagonal errors across all classes. This confirms that integrating spatial feature extraction with temporal dependency modeling enhances class separability. Overall, while all models achieve multi-class classification, the CNN–LSTM architecture demonstrates the most balanced and robust performance.



(a) (b) (c) (d)

Fig. 12. Event-related EEG topographic maps for four motor imagery directions: (a) left, (b) right, (c) front, and (d) back recorded using a five-channel Emotiv Insight headse

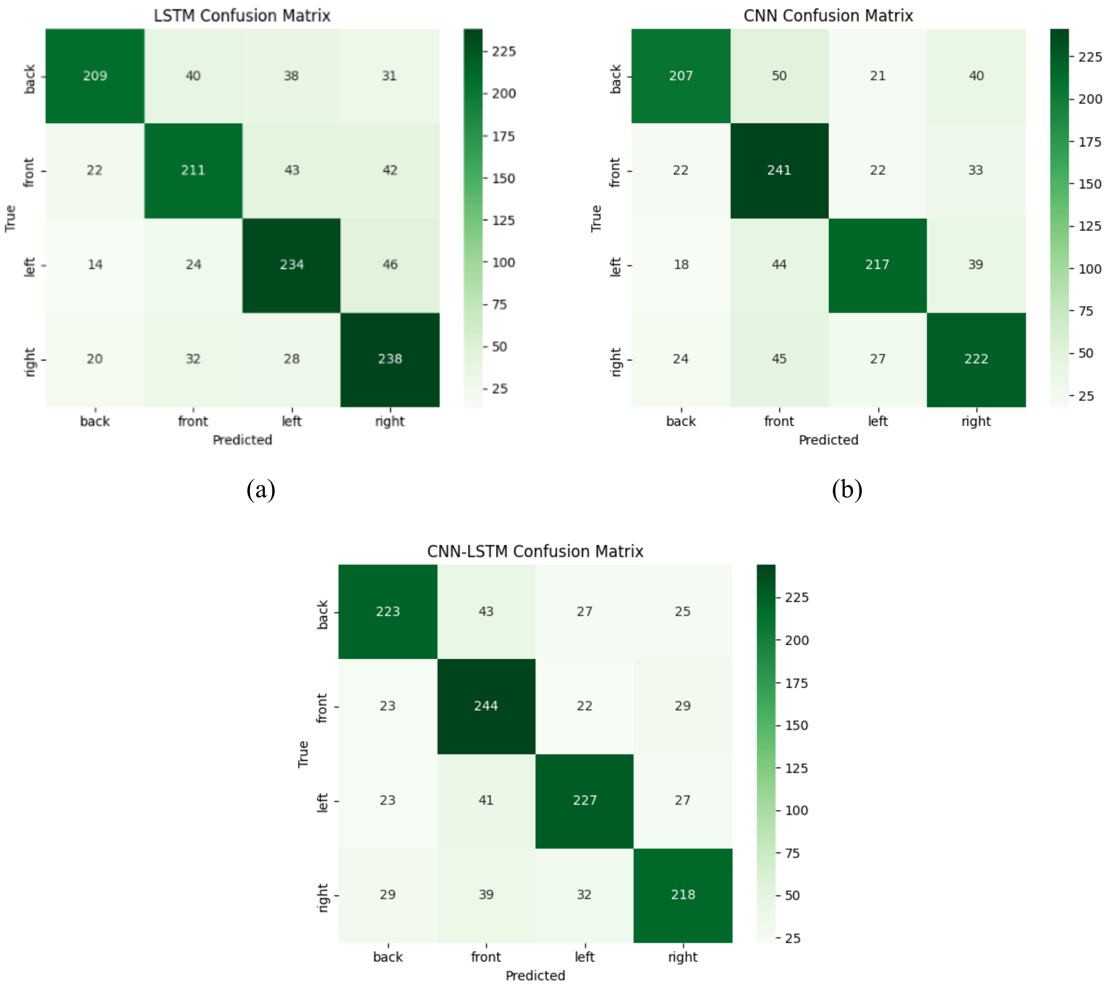
4.5. Model-Wise Classification Performance

Fig. 13 illustrate the relationships among window length, accuracy, latency, and Information Transfer Rate (ITR) for the LSTM, CNN, and hybrid CNN–LSTM models across different feature sets (2F–5F). Overall, accuracy increases consistently as window length increases for all models. The 4F and 5F feature sets generally achieve higher performance compared to 2F and 3F. However, longer windows also result in higher total latency, indicating a clear trade-off between classification performance and response time. The ITR curves show a similar upward trend with increasing window length, particularly beyond 0.7 s. Among the models, the hybrid CNN–LSTM achieves the highest overall accuracy and ITR, followed by CNN and LSTM. The CNN model demonstrates a favorable balance between accuracy and latency, while the LSTM model exhibits the highest latency growth at longer window sizes.

These results confirm that although longer windows improve classification accuracy, the optimal operating point should consider both latency and ITR, especially for real-time motor imagery applications. Paired t-tests between consecutive window sizes revealed that performance gains are not uniformly incremental but occur at specific transitions as shown in Table 3. The LSTM showed significant improvements ($p < 0.05$) primarily at larger windows (6→7 s, 10→11 s), while the CNN exhibited gains at mid-to-large transitions. The CNN–LSTM demonstrated the most consistent significant improvements at larger windows, confirming that joint spatial–temporal modeling benefits most from extended temporal context.

ITR analysis further supported these findings as shown in Table 3. Despite increased latency at longer windows, accuracy gains outweighed the latency cost across all models. The CNN–LSTM achieved the highest ITR (57.5 bits/min with 4 features), offering the best balance between decoding accuracy and communication efficiency, while the CNN remained competitive at lower latency and the LSTM yielded comparatively lower ITR.

As illustrated in Fig. 14, accuracy and ITR consistently increase with window length for all models, with the most pronounced gains beyond ≈ 7 s, particularly for the CNN–LSTM. These results confirm that extended temporal windows enhance classification robustness, and the hybrid architecture provides the most favorable accuracy–latency–ITR trade-off.



(a) (b)

(c)

Fig. 13. Confusion matrices of deep learning models for four-class motor imagery classification: (a) LSTM, (b) CNN, and (c) CNN–LSTM models evaluated on the test dataset

Table 3. Optimal window selection and computational trade-off analysis (Accuracy, ITR, and Latency)

Model	Feature Set	Best Window (s)	Accuracy	ITR (bits/min)	Latency (s)
LSTM	2F	11	0.6052	22.16	14.648
	3F	11	0.664	29.81	25.592
	4F	11	0.751	43.4	14.813
	5F	11	0.78	48.61	14.643
CNN	2F	11	0.6648	29.92	13.311
	3F	11	0.7204	38.29	12.503
	4F	11	0.7639	45.67	13.282
	5F	11	0.7341	40.53	12.627
CNN-LSTM	2F	11	0.7365	40.93	14.882
	3F	11	0.8058	53.57	13.627
	4F	11	0.8251	57.5	13.647
	5F	11	0.8163	55.67	14.868

Table 4, together with the confusion matrices, provides a comprehensive comparison of classification performance across CNN, LSTM, and CNN–LSTM models. Among the models tested, the CNN-LSTM network appeared to be more superior in terms of accuracy, with the highest average accuracy at 0.8714 and lowest standard deviation of 0.0927, in contrast to the CNN and LSTM networks. The reduced standard deviation shows improved generalization.

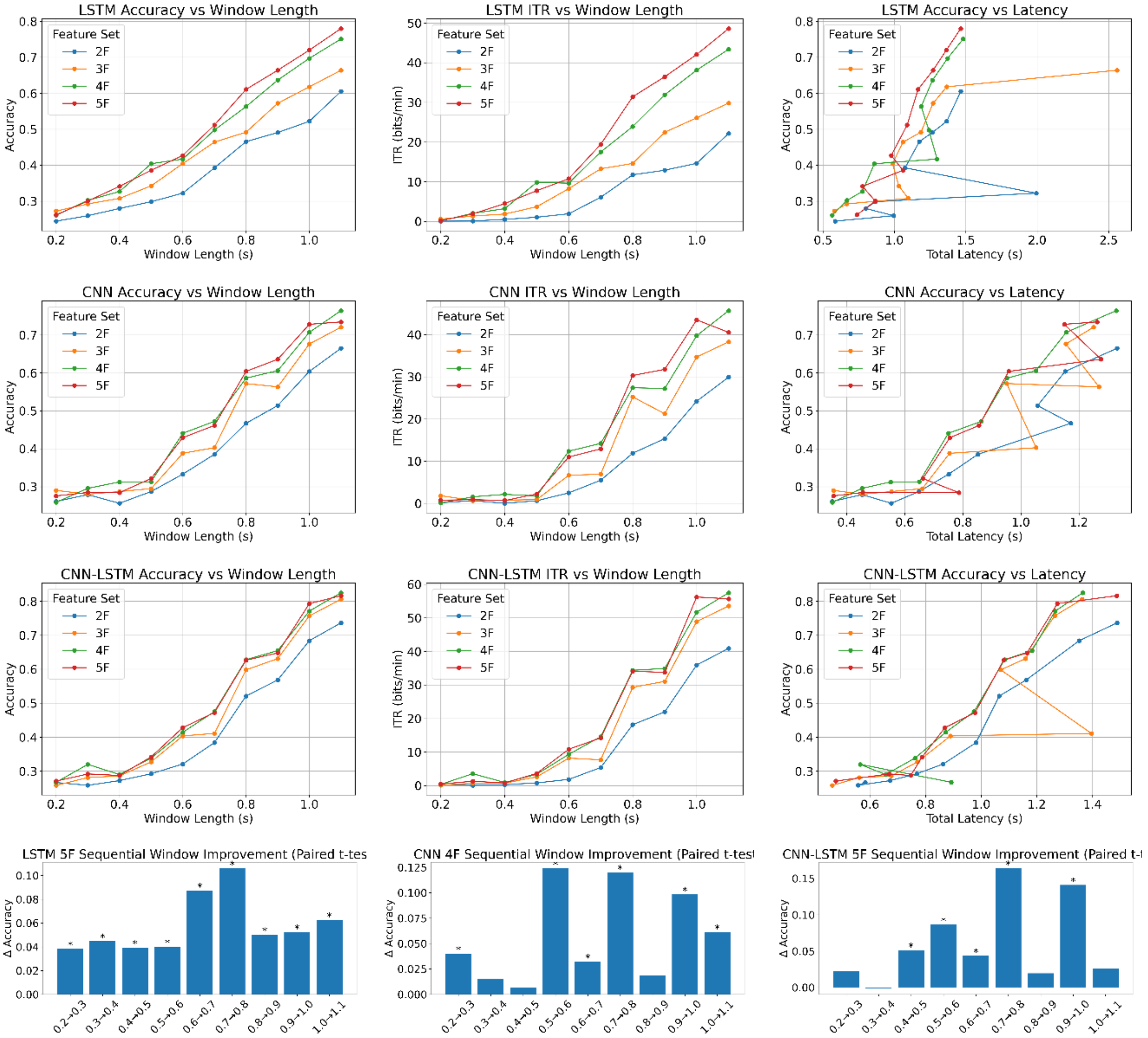


Fig. 14. Classification accuracy as a function of window size and feature set for different deep learning models

Learning the spatial information of the features effectively was evident in the CNN architecture, represented by its comparatively high mean accuracy of 0.8379; yet, the extent of accuracy variation from the lowest accuracy of 0.4073 may indicate its vulnerability to temporal changes and individual patterns within the data. On the other hand, the efficiency of the LSTM architecture in processing the temporal dependencies within the data was notable but showed the highest level of accuracy variation, with a standard deviation of 0.1822, when the spatial distinction among the EEG channels is restricted.

Analysis of confusion matrices also corroborates these results. The CNN-LSTM network reveals greater diagonal dominance, especially with respect to left and right motor imagery classes. This shows that class separation is improved. The confusion observed between front and back classes in some networks like CNN and LSTM is significantly diminished in this hybrid network. This implies that the spatial and temporal features captured jointly in the CNN-LSTM network allow for greater discriminative mapping of fine patterns of motor imagery.

The higher value of the minimum accuracy (0.5974) of the CNN-LSTM model further emphasizes the robustness of the model, which works perfectly even for the unfavorable setting. On the whole, the above results clearly illustrate that the combination of spatial, as well as temporal features, reinforces the accuracy of the classification, making the CNN-LSTM a most suitable model for multi-class motor imagery-EEG decoding for low-channel, consumer-grade EEG.

Table 4. Overall classification performance across models

Model	Mean Accuracy	Std. Dev.	Min Accuracy	Max Accuracy
CNN	0.8379	0.1170	0.4073	0.9697
LSTM	0.7974	0.1822	0.2884	0.9881
CNN–LSTM	0.8714	0.0927	0.5974	0.9826

4.6. Distribution of Subject-Level Performance

Fig. 15 is a box plot showing the overall subject-level mean classification accuracy for the proposed models of CNN, LSTM, and CNN-LSTM. It is clear that the CNN-LSTM model has the highest median accuracy with less variability, whereas the LSTM model has the widest box with lower minimum points, indicating that the LSTM is more susceptible to intra-subject variability and timing irregularities in the EEG signal.

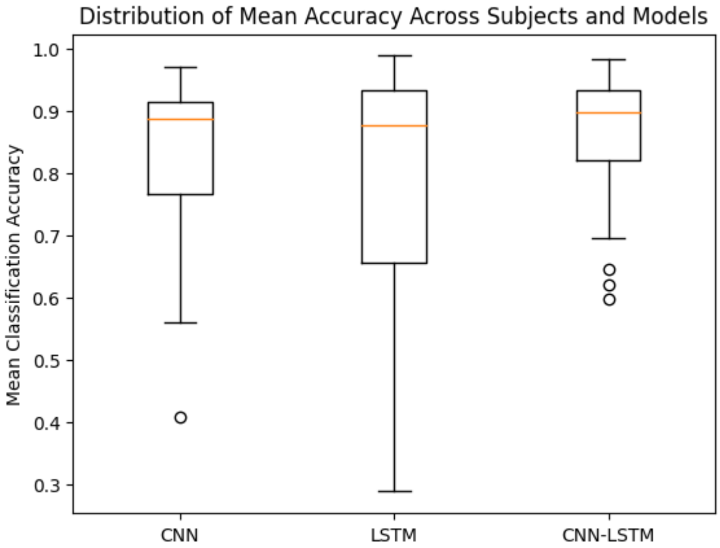


Fig. 15. Distribution of subject-wise mean classification accuracy across CNN, LSTM, and CNN–LSTM models

The CNN architecture portrays a moderate level of variability with high accuracy rates for numerous subjects but simultaneously presenting a number of low-performing subjects as outliers. The pattern revealed by this architecture shows that although the extraction of spatial features proves effective for certain subjects, generalization for the group as a whole might not be as effective.

From the graph of subject-level accuracy distributions, it is evident that the CNN-LSTM approach witnesses a marked increase in median value accuracy with a smaller interquartile range, which indicates better robustness with respect to subjects, while the LSTM approach has more variability and a larger accuracy range, which suggests its sensitive nature with respect to accuracy among subjects.

4.7. Class-Wise Performance Analysis

In order to further examine classification trend patterns, class accuracy scores were calculated and averaged over all subjects for each model. The corresponding data is presented in Table 5. Analysis of class accuracy scores shows that the CNN-LSTM approach is uniformly better than individual CNN and LSTM models for classification of motor imagery in all four directions. More specifically, the highest accuracy scores belong to back and left motor imagery classes (0.9006 and 0.8706 respectively), implying that these motor imagery directions correspond to more distinctive patterns in human brain activity. The accuracy is also found to be lowest for front class in all models, confirming its relatively diffuse and non-lateralized representation in human brain that is harder to discriminate. The performance of the CNN-LSTM approach is also significantly more well-rounded and accurate.

Table 5. Mean class-wise accuracy across subjects

Model	Back	Front	Left	Right
CNN	0.8819	0.7896	0.8627	0.8174
LSTM	0.8328	0.7829	0.7797	0.7942
CNN–LSTM	0.9006	0.8371	0.8706	0.8772

5. Discussion

This study demonstrates that multi-directional motor imagery decoding is feasible using a low-channel consumer-grade EEG device when appropriate spatial–temporal modeling is applied. Despite the limited spatial resolution of the five-channel Emotiv Insight headset, discriminative neural patterns were successfully extracted. Across all experiments, the hybrid CNN–LSTM architecture consistently achieved superior accuracy, lower variability, and higher robustness compared to standalone CNN and LSTM models. The results further confirm that temporal window length plays a critical role, with extended temporal integration significantly enhancing classification stability.

5.1. Impact of Low-Channel EEG and Signal Conditioning

A key issue with in this research is the limited spatial information captured using the five-channel Emotiv Insight EEG headset. Analysis using ICA showed that the good spatial features and the absence of prominent ocular and muscle artifacts in the retained components indicate that the essential neural information has been maintained despite the limited number of channels and the associated spatial constraints. The co-existence of the posterior-parietal components linked to visuospatial processing indicates the dominance of the video-based stimulus paradigm used in the experiment. Compared to prior studies employing high-density EEG systems (e.g., BCI Competition datasets), the achieved accuracy is slightly lower, which is expected due to reduced spatial resolution. However, when compared to recent works using consumer-grade EEG devices, the proposed approach achieves competitive or improved performance while maintaining lower computational complexity. Unlike transformer-based or multi-branch deep models reported in SOTA literature, the present architecture remains lightweight and suitable for edge deployment.

5.2. Neurophysiological Interpretation of Event-Related Topographies

The topographic mappings of event-related activity revealed class-specific and reproducible mappings, even with the rather low spatial resolution of the headset. The lateral motor imagery conditions (left and right motor imagery) revealed complementary hemispheric patterns, as commonly reported in MI studies, while front and back-related activity revealed more distributed patterns in the anterior and posterior regions of activation, likely due to higher cognitive loads and lack of specificity in those conditions. These results give a neurophysiological explanation for why there has been a constant decrease in classification accuracy for the front class in all models tested. Crucially, it is important to note that, despite not allowing for precise cortical localizations, it has been shown that direction-dependent activity patterns could be recovered even for low-resolution EEG data.

5.3. Advantages of Spatial–Temporal Feature Integration

Model-wise performance analysis clearly shows that architectures using purely spatial, CNN, or temporal, LSTM, modeling are insufficient to model the complexity of multi-directional motor imagery in low-channel settings. Although the CNN model effectively exploited the spatial relationships between the channels, it showed sensitivity regarding temporal variability and subject-dependent dynamics. In contrast, the LSTM model captured temporal dependencies but showed difficulty in distinguishing classes that were similar in a spatial sense, resulting in higher dispersion in performances across subjects.

The hybrid CNN–LSTM architecture addressed these shortcomings by using spatial feature extraction together with temporal sequence modeling. Consequently, class separability increased and misclassification between over-lapping classes was reduced, especially for front and back, with a variance in performance at the subject level. The decreased standard deviation and higher minimum accuracy seen for CNN–LSTM show increased robustness and generalization. These are two critical needs for BCI to function well in the real world.

5.4. Effect of Temporal Windowing and Feature Dimensionality

Compared to recent state-of-the-art deep architectures employing multi-layer CNN stacks, attention mechanisms, or transformer-based designs, the proposed CNN–LSTM model is deliberately lightweight. It consists of only two convolutional layers and a single LSTM layer, resulting in substantially lower computational complexity and reduced parameter count. This architectural simplicity makes the model more suitable for edge deployment, particularly in assistive BCI applications where computational resources and power consumption are limited. The measured inference times further confirm that the model can operate within practical real-time constraints.

The window size analysis for dimensionality of features showed that for all models, an increase in window size resulted in better classification performance. This observation highlights the significance of incorporating high-quality long-term information in MI-related EEG data. The performance gain for CNN standalone models plateaued after a certain window size, whereas for LSTM and CNN-LSTM models, there was a continued improvement due to the capability to learn from the temporal dependencies present in the data.

In addition, the more features were combined (number of channels), the more accuracy improve in both models. This confirms the idea mentioned earlier in the introduction, because with a minimal addition to the spatial information, a substantial boost in decoding accuracy can be achieved. From the study, to balance the window size and dimension of features, having a fix window size with a minimum dimension of features or vice versa could be optimal.

5.5. Subject-Level Variability and Generalization

The analysis of performance at the subject level confirmed the significant inter-subject variability, which had been known to be one of the major challenges in MI-based BCIs. The median accuracy was higher, and the interquartile range was narrower for the CNN-LSTM model compared to CNN and LSTM models, indicating better consistency across users. The LSTM model had the widest distribution in accuracy, indicating its sensitivity to neural dynamics that could vary largely among individuals.

These findings suggest that joint spatial-temporal learning benefits from increased diversity in a dataset, and this improved robustness is a feature of multi-subject training for the CNN-LSTM architecture. In particular, scalable BCI systems, where subject-specific calibration is not feasible, can be designed by keeping this finding in mind. Overall, generalized MI decoding using consumer-grade EEG is possible provided sufficient subject diversity and structured temporal representations form part of the learning pipeline.

5.6. Implications for Practical BCI Systems

In application terms, the results of the present work confirm the feasibility of using MI-based BCIs with lightweight, off-the-shelf EEG electrodes. With the proposed method, the reduction in the complexity of the experiment and comfort level, while preserving competitive accuracy, overcomes some of the difficulties of using the method in practice. It seems that high-density electrodes are not necessarily required for multi-directional decoding in the experiment based on the efficiency of deep techniques in the real experiment.

5.7. Limitations and Future Directions

Although encouraging outcomes have been achieved, a number of limitations still deserve consideration. On the one hand, a fixed low-channel setup will naturally impair spatial resolution, preventing delicate interpretations related to neurophysiological processes. On the other hand, although inter-subject learning has promoted generalization, learning within a subject-independent scenario is still a pending issue, especially for people with non-standard brain patterns. Future research will tap into transfer learning, domain adaptation, and attention-driven architectures to enhance learning capabilities. Closed-loop BCI learning and real-time learning will also be crucial for testing learning usability.

6. Conclusion

This study systematically demonstrates that multi-directional motor imagery can be reliably decoded using low-channel consumer-grade EEG by jointly optimizing temporal window length and spatial–temporal deep learning architectures. Using data from 53 participants recorded with the Emotiv Insight headset, we showed that extended temporal integration significantly improves classification performance, and that the hybrid CNN–LSTM architecture consistently outperforms standalone CNN and LSTM models. The results confirm that integrating convolutional spatial feature extraction with recurrent temporal modeling enhances robustness under limited spatial resolution conditions.

From a theoretical perspective, the findings highlight the critical role of temporal context in compensating for reduced channel density in EEG-based motor imagery decoding. The observed performance gains beyond mid-range window lengths indicate the presence of a temporal integration threshold necessary for stable discrimination of directional motor imagery patterns in low-channel systems.

Regarding real-time applicability, the total system latency comprises the temporal window duration and model inference time, the latter estimated at approximately 15 s for the CNN-LSTM architecture. For the evaluated window sizes, this latency is within a range suitable for low-speed assistive device control, such as wheelchair navigation or robotic arm positioning, though it remains insufficient for applications requiring. Real-time deployment will require online artifact removal and adaptive recalibration, which are identified as key future directions.

Although the proposed framework achieves competitive accuracy and favorable information transfer rates, several limitations remain. The study relies on controlled experimental conditions and a five-channel configuration, which may restrict generalization to fully subject-independent or real-world dynamic environments. In addition, online artifact handling and adaptive recalibration were not implemented.

Future work will focus on lightweight edge deployment optimization, and real-time adaptive calibration strategies to enhance robustness under practical operating conditions. These directions are essential for translating low-channel EEG-based motor imagery systems from laboratory validation to reliable assistive BCI applications.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding: This research received no external funding.

Acknowledgment: The authors would like to express their sincere gratitude to Rajamangala University of Technology Thanyaburi (RMUTT), Thailand, for providing financial support through a Ph.D. scholarship. The authors also thank the participants who voluntarily took part in the EEG data collection and contributed to this study. Special appreciation is extended to colleagues for their technical support and constructive discussions throughout the research process.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] A. Saibene, M. Caglioni, S. Corchs, and F. Gasparini, “EEG-Based BCIs on Motor Imagery Paradigm Using Wearable Technologies: A Systematic Review,” *Sensors*, vol. 23, no. 5, p. 2798, 2023, https://doi.org/10.3390/s23052798.

[2] X. Y. Liu *et al.*, “Recent applications of EEG-based brain-computer-interface in the medical field,” *Military Medical Research*, vol. 12, no. 14, 2025, https://doi.org/10.1186/s40779-025-00598-z.

[3] K. Värbu, N. Muhammad, and Y. Muhammad, “Past, present, and future of EEG-based BCI applications,” *Sensors*, vol. 22, no. 9, p. 3331, 2022, https://doi.org/10.3390/s22093331.

[4] H. Hu *et al.*, "A Survey on Brain-Computer Interface-Inspired Communications: Opportunities and Challenges," *IEEE Communications Surveys & Tutorials*, vol. 27, no. 1, pp. 108-139, 2025, https://doi.org/10.1109/COMST.2024.3396847.

[5] C. He *et al.*, "Diversity and Suitability of the State-of-the-Art Wearable and Wireless EEG Systems Review," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 8, pp. 3830-3843, 2023, https://doi.org/10.1109/JBHI.2023.3239053.

[6] M. A. Khan, R. Das, H. K. Iversen, and S. Puthusserypady, “Review on motor imagery based BCI systems for upper limb post-stroke neurorehabilitation: From designing to application,” *Computers in Biology and Medicine*, vol. 123, p. 103843, 2020, https://doi.org/10.1016/j.compbiomed.2020.103843.

[7] A. E. Gentile *et al.*, “Motor imagery for paediatric neurorehabilitation: how much do we know? Perspectives from a systematic review,” *Frontiers in Human Neuroscience*, vol. 18, p. 1245707, 2024, https://doi.org/10.3389/fnhum.2024.1245707.

[8] A. Palumbo, V. Gramigna, B. Calabrese, and N. Ielpo, “Motor-imagery EEG-based BCIs in wheelchair movement and control: A systematic literature review,” *Sensors*, vol. 21, no. 18, p. 6285, 2021, https://doi.org/10.3390/s21186285.

[9] P. Tantawanich, C. Phunruangsakao, S. -I. Izumi and M. Hayashibe, "A Systematic Review of Bimanual Motor Coordination in Brain-Computer Interface," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 33, pp. 266-285, 2025, https://doi.org/10.1109/TNSRE.2024.3522168.

[10] H. Yadav and S. Maini, “Decoding brain signals: A comprehensive review of EEG-based BCI paradigms, signal processing and applications,” *Computers in Biology and Medicine*, vol. 196, p. 110937, 2025, https://doi.org/10.1016/j.compbiomed.2025.110937.

[11] R. Sharma and H. K. Meena, “Emerging trends in EEG signal processing: A systematic review,” *SN Computer Science*, vol. 5, no. 4, p. 415, 2024, https://doi.org/10.1007/s42979-024-02773-w.

[12] Y. E. Lee, G. H. Shin, M. Lee, and S. W. Lee, “Mobile BCI dataset of scalp-and ear-EEGs with ERP and SSVEP paradigms while standing, walking, and running,” *Scientific Data*, vol. 8, no. 1, p. 315, 2021, https://doi.org/10.1038/s41597-021-01094-4.

[13] H. Guillen-Sanz, D. Checa, I. Miguel-Alonso, and A. Bustillo, “A systematic review of wearable biosensor usage in immersive virtual reality experiences,” *Virtual Reality*, vol. 28, no. 2, p. 74, 2024, https://doi.org/10.1007/s10055-024-00970-9.

[14] L. F. H. Contreras *et al.*, “Neuromorphic neuromodulation: Towards the next generation of closed-loop neurostimulation,” *PNAS Nexus*, vol. 3, no. 11, p. 488, 2024, https://doi.org/10.1093/pnasnexus/pgae488.

[15] Q. Liu, L. Yang, Z. Zhang, H. Yang, Y. Zhang, and J. Wu, “The feature, performance, and prospect of advanced electrodes for electroencephalogram,” *Biosensors*, vol. 13, no. 1, p. 101, 2023, https://doi.org/10.3390/bios13010101.

[16] D. Gorjan, K. Gramann, K. De Pauw, and U. Marusic, “Removal of movement-induced EEG artifacts: current state of the art and guidelines,” *Journal of Neural Engineering*, vol. 19, no. 1, p. 011004, 2022, https://doi.org/10.1088/1741-2552/ac542c.

[17] P. Sawangjai, S. Hompoonsup, P. Leelaarporn, S. Kongwudhikunakorn and T. Wilaiprasitporn, "Consumer Grade EEG Measuring Sensors as Research Tools: A Review," *IEEE Sensors Journal*, vol. 20, no. 8, pp. 3996-4024, 2020, https://doi.org/10.1109/JSEN.2019.2962874.

[18] F. Ishtiaque *et al*., “Systematic comparison between a research-grade EEG device and a consumer-grade BCI device for predicting consumer preference using an ML framework,” *Multimedia Tools and Applications*, vol. 84, no. 34, pp. 42807–42825, 2025, https://doi.org/10.1007/s11042-025-20812-3.

[19] A. Al-Saegh, S. A. Dawwd, and J. M. Abdul-Jabbar, “Deep learning for motor imagery EEG-based classification: A review,” *Biomedical Signal Processing and Control*, vol. 63, p. 102172, 2021, https://doi.org/10.1016/j.bspc.2020.102172.

[20] Y. Li *et al*., "A Tale of Single-Channel Electroencephalography: Devices, Datasets, Signal Processing, Applications, and Future Directions," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1-20, 2025, https://doi.org/10.1109/TIM.2025.3556900.

[21] F. M. Garcia-Moreno, M. Bermudez-Edo, J. L. Garrido, and M. J. Rodríguez-Fórtiz, “Reducing Response Time in Motor Imagery Using A Headband and Deep Learning,” *Sensors*, vol. 20, no. 23, p. 6730, 2020, https://doi.org/10.1007/s42979-024-02773-w.

[22] V. Peterson, C. Galván, H. Hernández, and R. Spies, “A feasibility study of a complete low-cost consumer-grade brain-computer interface system,” *Heliyon*, vol. 6, no. 3, p. e03425, 2020, https://doi.org/10.1016/j.heliyon.2020.e03425.

[23] G. Prapas, K. Glavas, K. D. Tzimourta, A. T. Tzallas, and M. G. Tsipouras, “Mind the Move: Developing a Brain-Computer Interface Game with Left-Right Motor Imagery †,” *Information*, vol. 14, no. 7, p. 354, 2023, https://doi.org/10.3390/info14070354.

[24] P. Tarara, I. Przybył, J. Schöning, and A. Gunia, “Motor imagery-based brain-computer interfaces: an exploration of multiclass motor imagery-based control for Emotiv EPOC X,” *Frontiers in Neuroinformatics*, vol. 19, 2025, https://doi.org/10.3389/fninf.2025.1625279.

[25] S. Akuthota *et al*., “Enhancing Real-Time Cursor Control with Motor Imagery and Deep Neural Networks for Brain–Computer Interfaces,” *Information*, vol. 15, no. 11, p. 702, 2024, https://doi.org/10.3390/info15110702.

[26] J. Sabio, N. S. Williams, G. M. McArthur, and N. A. Badcock, “A scoping review on the use of consumer-grade EEG devices for research,” *PLoS One*, vol. 19, no. 3, p. e0291186, 2024, https://doi.org/10.1371/journal.pone.0291186.

[27] M. F. Hinss, E. S. Jahanpour, B. Somon, L. Pluchon, F. Dehais, and R. N. Roy, “Open multi-session and multi-task EEG cognitive Dataset for passive brain-computer Interface Applications,” *Scientific Data*, vol. 10, no. 1, p. 85, 2023, https://doi.org/10.1038/s41597-022-01898-y.

[28] H. B. Han, B. Kim, Y. Kim, Y. Jeong, and J. H. Choi, “Nine-day continuous recording of EEG and 2-hour of high-density EEG under chronic sleep restriction in mice,” *Scientific Data*, vol. 9, no. 1, p. 225, 2022, https://doi.org/10.1038/s41597-022-01354-x.

[29] V. Asanza, L. L. Lorente-Leyva, D. H. Peluffo-Ordóñez, D. Montoya, and K. Gonzalez, “MILimbEEG: A dataset of EEG signals related to upper and lower limb execution of motor and motor imagery tasks,” *Data in Brief*, vol. 50, p. 109540, 2023, https://doi.org/10.1016/j.dib.2023.109540.

[30] P. V. Dhole, D. G. Chaudhary, V. D. Dhangar, S. D. Shejul, S. S. Datwase, and B. W. Gawali, “A Review of EEG Artifact Removal Techniques for Brain-Computer Interface,” *SN Computer Science*, vol. 6, p. 1026, 2025, https://doi.org/10.1007/s42979-025-04592-z.

[31] A. Nogales and Á. J. García-Tejedor, "A Systematic Review of Electroencephalography Open Datasets and Their Usage With Deep Learning Models," *IEEE Access*, vol. 11, pp. 72384-72399, 2023, https://doi.org/10.1109/ACCESS.2023.3293421.

[32] J. Yang, L. Wang, W. Cai, L. Zhang, C. Xie, and Z. Wang, "EDANet: Efficient domain-adaptive attention neural network for EEG classification of motor imagery," *Expert Systems with Applications*, vol. 294, p. 128783, 2025, https://doi.org/10.1016/j.eswa.2025.128783.

[33] X. Geng, D. Li, H. Chen, P. Yu, H. Yan, and M. Yue, "An improved feature extraction algorithms of EEG signals based on motor imagery brain-computer interface," *Alexandria Engineering Journal*, vol. 61, no. 6, pp. 4807–4820, 2022, https://doi.org/10.1016/j.aej.2021.10.034.

[34] Y. Liang, M. Meng, Y. Gao, and X. Xi, "MVC-former adaptation: A multi-view convolution transformer-based domain adaptation framework for cross-subject motor imagery classification," *Neurocomputing*, vol. 649, p. 130875, 2025, https://doi.org/10.1016/j.neucom.2025.130875.

[35] H. Altaheri *et al.*, "Deep learning techniques for classification of electroencephalogram (EEG) motor imagery (MI) signals: a review," *Neural Computing and Applications*, vol. 35, no. 20, pp. 14681–14722, 2023, https://doi.org/10.1007/s00521-021-06352-5.

[36] J. Sabio, N. S. Williams, G. M. McArthur, and N. A. Badcock, "A scoping review on the use of consumer-grade EEG devices for research," *PLoS One*, vol. 19, no. 3, p. e0291186, 2024, https://doi.org/10.1371/journal.pone.0291186.

[37] R. Li, D. Yang, F. Fang, K. S. Hong, A. L. Reiss, and Y. Zhang, "Concurrent fNIRS and EEG for Brain Function Investigation: A Systematic, Methodology-Focused Review," *Sensors*, vol. 22, no. 15, p. 5865, 2022, https://doi.org/10.3390/s22155865.

[38] V. Peterson, C. Galván, H. Hernández, and R. Spies, "A feasibility study of a complete low-cost consumer-grade brain-computer interface system," *Heliyon*, vol. 6, no. 3, p. e03425, 2020, https://doi.org/10.1016/j.heliyon.2020.e03425.

[39] X. Wu, D. Long, and J. Yang, "Generative motor imagery dynamic networks: EEG-controlled grasping via individualized model training," *Cognitive Neurodynamics*, vol. 19, no. 1, p. 174, 2025, https://doi.org/10.1007/s11571-025-10360-9.

[40] B. Maiseli *et al.*, "Brain–computer interface: trend, challenges, and threats," *Brain Informatics*, vol. 10, p. 20, 2023, https://doi.org/10.1186/s40708-023-00199-3.

[41] A. Carretero and A. Araujo, "Design Decisions for Wearable EEG to Detect Motor Imagery Movements," *Sensors*, vol. 24, no. 15, p. 4763, 2024, https://doi.org/10.3390/s24154763.

[42] N. Sharma *et al.*, "Recent Trends in EEG-Based Motor Imagery Signal Analysis and Recognition: A Comprehensive Review," *IEEE Access*, vol. 11, pp. 80518-80542, 2023, https://doi.org/10.1109/ACCESS.2023.3299497.

[43] I. M. A. Wirawan *et al.*, "Acquisition and processing of Motor Imagery and Motor Execution Dataset (MIMED) for six movement activities," *Data in Brief*, vol. 56, p. 110833, 2024, https://doi.org/10.1016/j.dib.2024.110833.

[44] Z. Shuqfa, A. Lakas, and A. N. Belkacem, "Increasing accessibility to a large brain–computer interface dataset: Curation of physionet EEG motor movement/imagery dataset for decoding and classification," *Data in Brief*, vol. 54, p. 110181, 2024, https://doi.org/10.1016/j.dib.2024.110181.

[45] L. -L. Li, G. -Z. Cao, H. -J. Liang, Y. -P. Zhang and F. Cui, "Human Lower Limb Motion Intention Recognition for Exoskeletons: A Review," *IEEE Sensors Journal*, vol. 23, no. 24, pp. 30007-30036, 2023, https://doi.org/10.1109/JSEN.2023.3328615.

[46] W. Liao, J. Li, X. Zhang, and C. Li, "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*, vol. 17, 2023, https://doi.org/10.3389/fnhum.2023.1117670.

[47] X. Geng, D. Li, H. Chen, P. Yu, H. Yan, and M. Yue, "An improved feature extraction algorithms of EEG signals based on motor imagery brain-computer interface," *Alexandria Engineering Journal*, vol. 61, no. 6, pp. 4807–4820, 2022, https://doi.org/10.1016/j.aej.2021.10.034.

[48] N. Mateen *et al.*, "Subject based feature selection for hybrid brain computer interface using genetic algorithm and support vector machine," *Results in Engineering*, vol. 27, p. 105649, 2025, https://doi.org/10.1016/j.rineng.2025.105649.

[49] P. Biswas, S. Mondal, D. Chowdhury, S. Das and M. Laha, "BrainDepth: Brain Signal Analysis for Neuro-Fuzzy Modeling of Subjective Depth Perception," *2025 IEEE 3rd International Symposium on Sustainable Energy, Signal Processing and Cybersecurity*, pp. 1-6, 2025, https://doi.org/10.1109/iSSSC66652.2025.11388220.

[50] J. W. Choi *et al.*, "Neural Applications Using Immersive Virtual Reality: A Review on EEG Studies," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 31, pp. 1645-1658, 2023, https://doi.org/10.1109/TNSRE.2023.3254551.

[51] V. Peterson, C. Galván, H. Hernández, and R. Spies, "A motor imagery vs. rest dataset with low-cost consumer grade EEG hardware," *Heliyon*, vol. 6, no. 3, p. e03425, 2020, https://doi.org/10.1016/j.heliyon.2020.e03425.

[52] Y. Liang, M. Meng, Y. Gao, and X. Xi, "MVC-former adaptation: A multi-view convolution transformer-based domain adaptation framework for cross-subject motor imagery classification," *Neurocomputing*, vol. 649, p. 130875, 2025, https://doi.org/10.1016/j.neucom.2025.130875.

[53] J. Yang, L. Wang, W. Cai, L. Zhang, C. Xie, and Z. Wang, "EDANet: Efficient domain-adaptive attention neural network for EEG classification of motor imagery," *Expert Systems with Applications*, vol. 294, p. 128783, 2025, https://doi.org/10.1016/j.eswa.2025.128783.

[54] M. Spanos *et al.*, "Post Hoc Event-Related Potential Analysis of Kinesthetic Motor Imagery-Based Brain-Computer Interface Control of Anthropomorphic Robotic Arms," *Electronics*, vol. 14, no. 15, p. 3106, 2025, https://doi.org/10.3390/electronics14153106.

[55] G. Prapas, K. Glavas, K. D. Tzimourta, A. T. Tzallas, and M. G. Tsipouras, "Mind the Move: Developing a Brain-Computer Interface Game with Left-Right Motor Imagery †," *Information*, vol. 14, no. 7, p. 354, 2023, https://doi.org/10.3390/info14070354.

[56] J. Tong *et al.*, "Towards Improving Motor Imagery Brain–Computer Interface Using Multimodal Speech Imagery," *Journal of Medical and Biological Engineering*, vol. 43, no. 3, pp. 216–226, 2023, https://doi.org/10.1007/s40846-023-00798-9.