*Emerging Science Innovation, vol. 6, 2026, pp. 32-49*

# A Lightweight Real-Time EEG-Based Brain–Computer Interface for Imagined Speech-Controlled Robotics via Machine Learning

Hanif Adedotun[1,*], Precious Akah, Nyangwarimam Obadiah Ali

Department of Computer Engineering; Intelligent Systems Integration & Communications Research Group, Computer Engineering Department, Nile University of Nigeria, FCT, Nigeria

Received 13 December 2025; received in revised form 26 January 2026; accepted 28 January 2026

DOI: https://doi.org/10.46604/emsi.2026.15979

## Abstract

This study aims to develop a lightweight, real-time brain–computer interface system for classifying imagined speech commands using electroencephalography (EEG) signals acquired from a 14-channel Emotiv headset. Four motor-related imagined speech commands—push, pull, left, and right—are recorded from 15 participants, with a neutral state serving as a baseline. Continuous wavelet transform (CWT) is applied to extract time–frequency features from the EEG signals, which are converted into scalograms representations and used as inputs for machine learning models. Among the evaluated classifiers, XGBoost achieves the highest classification accuracy of 98.10%. The system is deployed on a Raspberry Pi-controlled robotic platform, enabling real-time inference within 900 ms. The results demonstrate efficiency, low computational overhead, and highlight the potential of consumer-grade EEG devices for real-time assistive robotics applications.

**Keywords:** brain–computer interface, electroencephalography (EEG), real-time imagined speech, XGBoost, assistive robotics

## 1. Introduction

Brain–computer interface (BCI) technology enables direct communication between the human brain and external devices without relying on the peripheral nervous system or muscular control. Over recent decades, significant advances in BCI research have focused on decoding neural signals to enable command generation based on imagined speech or motor intentions. Among various signal acquisition modalities, including functional magnetic resonance imaging (fMRI), electrocorticography (ECoG), and electroencephalography (EEG), EEG remains the most widely adopted approach due to its non-invasive nature, low cost, and excellent temporal resolution [1-3].

EEG-based BCI systems have been extensively investigated for decoding imagined speech and motor-related commands. During these tasks, distinct neural activity patterns are captured through scalp-mounted electrodes, enabling the classification of user intentions [2-3]. Although promising performance has been reported in controlled laboratory environments [1], challenges remain in achieving real-time deployment on low-power embedded platforms.

Shashidhar et al. developed an EEG-based Brain-Sense mobility system implemented on a Raspberry Pi using Python for signal processing [1]. Chaurasia et al. introduced Brain-Bot for unmanned ground vehicle (UGV) control through EEG and GPS fusion, although the system is constrained by a limited dataset [2]. Asanza et al. implemented steady-state visual evoked potential (SSVEP)-based EEG classification using support vector machine (SVM), multilayer perceptron (MLP), random

1* Corresponding author. E-mail address: hanif.adedotun@gmail.com

forest (RF), K-nearest neighbors (KNN), and XGBoost, achieving moderate accuracy levels [4]. Similarly, Rakhmatulin et al. utilized the fast fourier transform for signal processing and developed a low-cost Raspberry Pi-based BCI robot controller [3]. Lin et al. adopted principal components analysis-based feature extraction for EEG-based quadcopter navigation using the Emotiv EPOC headset [5].

Recent studies further explored deep learning techniques for imagined speech decoding. Das et al. achieved 96% accuracy using convolutional neural networks (CNNs) models with 14-channel EEG input [6], while Saha et al. proposed a hierarchical CNN system for imagined speech recognition, yielding 83.42% accuracy [7]. However, Kumar et al. reported only 23.7% accuracy using a CNN-Long Short-Term Memory (LSTM) hybrid model, primarily limited by dataset size [8].

Broader challenges related to feature generalization and domain integration were discussed in a review by Hossain et al. [9], and further addressed by Wang et al. through a deep feature learning framework using target priors in an ECoG-based BCI system [10]. Chaudhry et al. integrated EEG and Arduino to control a prosthetic arm via the Emotiv EPOC+ headset [11], and Zgallai et al. developed a smart wheelchair control system using EEG-based BCI input [12].

Despite technological advances, non-invasive BCI systems continue to face challenges in signal variability, headset limitations, and inconsistent preprocessing protocols [13]. These limitations reduce classification robustness and hinder real-world deployment. To address these issues, researchers have emphasized the need for improved signal processing pipelines and machine learning algorithms tailored to EEG signal dynamics [14].

In response, this study proposes a real-time imagined speech BCI system using a 14-channel Emotiv Epoc X headset and a Raspberry Pi 4 controller. The system allows participants to control a four-wheeled robotic platform through four imagined speech commands: push, pull, left, and right. EEG signals are preprocessed using independent component analysis (ICA) and bandpass filtering. Time–frequency features are extracted using continuous wavelet transform (CWT), while classification is performed using an XGBoost model trained on scalogram representations.

This work aims to expand BCI accessibility by using affordable, consumer-grade EEG equipment and focusing on real-time integration. Data are collected from 15 volunteers with diverse hair types to enhance robustness and extend prior work. EEG electrodes are positioned according to brain regions responsible for motor planning, speech production, and auditory processing.

The proposed system integrates signal acquisition, processing, classification, and robotic actuation into a single pipeline. It also supports real-time feedback and scalability for future emotion recognition modules [15].

The main contributions of this paper are summarized as follows:

(1) A wireless EEG acquisition approach using a 14-channel emotiv headset.

(2) Application of CWT for time–frequency feature extraction.

(3) Implementation of XGBoost for classification, benchmarked against six other classifiers.

(4) Integration of imagined speech classification with real-time robotic control under 900ms.

(5) Evaluation using a robust dataset comprising 15 subjects and five command classes (four actions and a neutral state).

The remaining part of the paper is organized as follows: Section II presents data acquisition and methodology, including preprocessing and model training. Section III details classification results and performance evaluation. Section IV discusses conclusions and outlines directions for future work.

## 2. Creative Mechanism Design Methodology

This study presents the design and implementation of a non-invasive BCI system capable of translating EEG signals into actionable commands for robotic control. The overall workflow of the proposed system is shown in Fig. 1, including EEG acquisition (14-channel Emotiv Epoc X), signal preprocessing (bandpass filtering and ICA), feature extraction using CWT, classification via XGBoost, and real-time robotic control using Raspberry PI.

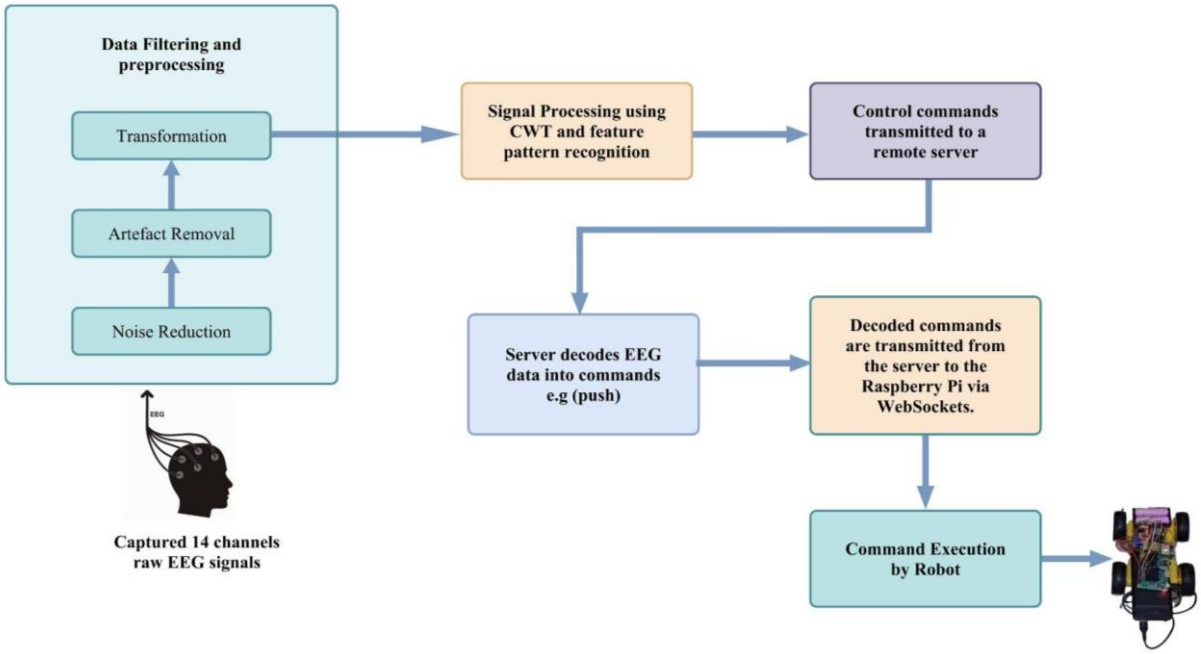



Fig. 1 The block diagram of the proposed BCI system

The pipeline begins with EEG signal acquisition using a 14-channel Emotiv Epoc X headset. Acquired signals undergo preprocessing, which includes bandpass filtering (8–30 Hz) and artifact removal using ICA. After preprocessing, time–frequency features are extracted using CWT to generate high-resolution scalograms.

These extracted features are then used as inputs for training multiple machine learning models. Among them, XGBoost is selected based on its superior accuracy and computational efficiency. During real-time operation, the EEG headset continuously streams data to a local machine for preprocessing and classification. The predicted command is transmitted to a Raspberry Pi-controlled robotic platform via the WebSocket protocol.

The Raspberry Pi interprets the incoming command and activates corresponding motor controls using an L298N driver circuit. This architecture enables seamless real-time communication between brain signals and robotic actions, supporting hands-free control for assistive applications.

### *2.1. EEG signal acquisition and preprocessing*

Data are obtained using a 14-channel Emotiv Epoc X EEG headset. Electrodes are positioned according to the international 10–20 system at the following sites: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. Electrodes AF3, AF4, F3, F4, F7, and F8 capture neuronal activity in the frontal lobe (lobus frontalis), while FC5, FC6, T7, and T8 monitor the temporal lobe (lobus temporalis). P7 and P8 electrodes target the parietal lobe (lobus parietalis), and O1 and O2 record activity from the occipital lobe (lobus occipitalis), as illustrated in Fig. 2.

The acquired signals undergo preprocessing, including bandpass filtering (8–30 Hz) to isolate relevant frequency bands and ICA to remove ocular and muscular artifacts. CWT is applied to extract high-resolution time–frequency features, represented as scalograms.

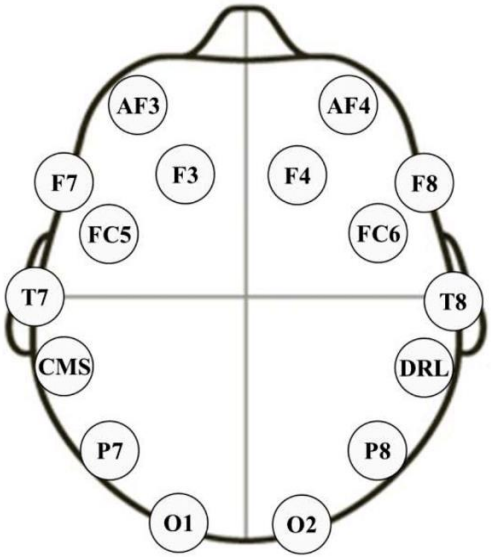



Fig. 2 Electrode layout of the 14-channel Emotiv EPOC+ EEG headset

*2.2. Experimental protocol for inner speech EEG acquisition*

To collect EEG data corresponding to inner speech, each experimental trial is structured into distinct phases to ensure clean segmentation of cognitive states. As illustrated in Fig. 3, each trial begins with a 5-second query phase, during which a subject is presented with a word or command, either visually or auditorily. This is followed by a 5-second preparation period, allowing the participant to cognitively process the prompt without initiating any mental speech. The critical inner speech phase lasts 8 seconds, during which the participant is instructed to silently and repeatedly imagine articulating the given word or phrase without any actual vocalization or muscle movement.

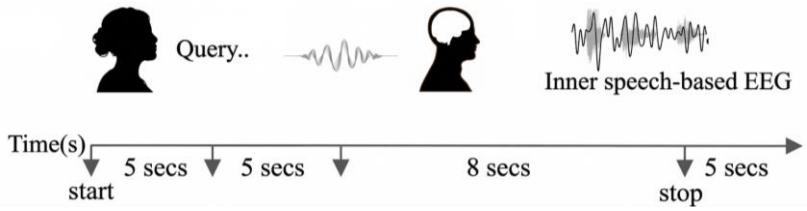



Fig. 3 Inner-speech EEG trial timeline

EEG data recorded during this phase is later used for feature extraction and classification. Finally, each trial concludes with a 5-second rest period, enabling the brain to return to a baseline state and minimizing carry-over effects into the next trial. This well-structured protocol is designed to isolate neural activity related specifically to imagined speech, ensuring temporal consistency and data quality for subsequent machine learning tasks.

*2.3. Feature matrix construction*

The extracted EEG dataset is structured into a feature matrix comprising 2,132 samples and 2,549 features, as shown in Table 1. Each row represents a distinct temporal segment or epoch derived from the preprocessed EEG recordings. The columns correspond to statistical and signal-derived metrics, such as mean values computed over specific EEG channels, frequency bands, or time-frequency components.

Feature names follow a consistent nomenclature (e.g., mean_0_a, mean_d_2_a), suggesting combinations of spatial (channel-based), spectral (frequency band), and temporal descriptors. This high-dimensional feature space captures complex patterns of neural activity associated with imagined speech. Prior to classification, the dataset is processed with dimensionality reduction, normalization, and alignment with ground truth labels to enhance model performance.

Table 1 Extracted data

| | mean_0_a | mean_1_a | mean_2_a | mean_3_a | mean_4_a | mean_d_0_a | mean_d_1_a | mean_d_2_a | mean_d_3_a | mean_d_4_a |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 4.620 | 30.3 | -356.0 | 15.60 | 26.3 | 1.070 | 0.411 | -15.700 | 2.060 | 3.15 |
| 1 | 28.800 | 33.1 | 32.0 | 25.80 | 22.8 | 6.550 | 1.680 | 2.880 | 3.830 | -4.82 |
| 2 | 8.900 | 29.4 | -416.0 | 16.70 | 23.7 | 79.900 | 3.360 | 90.200 | 89.900 | 2.03 |
| 3 | 14.900 | 31.6 | -143.0 | 19.80 | 24.3 | -0.584 | -0.294 | 8.620 | 2.300 | -1.97 |
| 4 | 28.300 | 31.3 | 45.2 | 27.30 | 24.5 | 34.800 | -5.790 | 3.060 | 41.400 | 5.52 |
| 2127 | 32.400 | 32.2 | 32.2 | 30.80 | 23.4 | 1.640 | -2.030 | 0.647 | -0.121 | -1.10 |
| 2128 | 16.300 | 31.3 | -284.0 | 14.30 | 23.9 | 4.200 | 1.090 | 4.460 | 4.720 | 6.63 |
| 2129 | -0.547 | 28.3 | -259.0 | 15.80 | 26.7 | 0.980 | 6.900 | 12.700 | 2.030 | 4.64 |
| 2130 | 16.800 | 19.9 | -288.0 | 8.34 | 26.0 | 2.430 | 1.580 | -16.000 | 1.690 | 4.74 |
| 2131 | 27.000 | 32.0 | 31.8 | 25.00 | 28.9 | 4.990 | 1.950 | 6.210 | 3.490 | -3.51 |

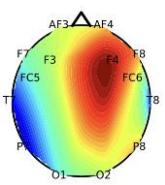
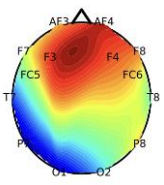
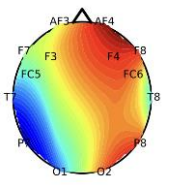
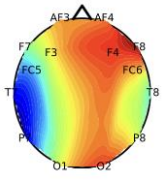
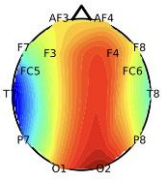
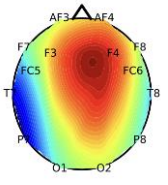
### *2.4. Data visualization using topographical maps*

EEG recordings capture electrical activity only at discrete electrode positions, necessitating spatial interpolation to estimate voltage distribution across the entire scalp. Topographical maps provide a smooth, continuous visual representation of brain activity using interpolation techniques such as spherical spline interpolation [15]. These maps offer intuitive insight into the spatial patterns of EEG data.

Topographical visualization complements spatial (bandpower) and temporal (event-related potential, ERP) analyses, providing spatial context that enables a comprehensive understanding of neuronal activity [16]. Bandpower distribution across various regions can reveal localized brain responses linked to imagined speech tasks.

In this study, topological data analysis (TDA) is used to examine variations in brain activity patterns under different imagined command conditions (e.g., imagining left vs. right hand movement). Table 2 displays the topographical maps illustrating average band power distributions across the scalp for each class of imagined speech activity.

Table 2 Topographical EEG maps for Sample Subject 1

| S/N | 1 | 2 | 3 |
|---|---|---|---|
| Description | Theta band (4--8 Hz) | Alpha band (8--12 Hz) | Beta band (12--30 Hz) |
| Eyes Open |  |  |  |
| Stress |  |  |  |

Topographic EEG analyses reveal distinct spectral and spatial patterns in theta (4–8 Hz) and alpha (8–12 Hz) bands under Eyes Open and Stress conditions. In the theta band, stress elicits a pronounced increase in frontal and central regions, particularly at F4, aligning with prior findings linking frontal midline theta to increased cognitive effort and emotional regulation. Conversely, alpha activity, primarily frontally distributed during the Eyes Open condition, shifts posteriorly and becomes more widespread under stress. This pattern suggests a functional disengagement or cortical inhibition; as posterior alpha increases are associated with reduced sensory processing. The contrasting spatial dynamics—frontal theta enhancement versus posterior alpha elevation—highlight the multidimensional neural impact of stress and underscore the importance of multiband analysis for robust BCI systems. These insights support integrating both theta and alpha features in real-time imagined speech decoding models, particularly in emotionally dynamic or mentally demanding environments.

The Emotiv Epoc X headset provides configurable parameters via its built-in software. In this study, the sampling rate is configured to 256 Hz, and the system is set to record from 14 active channels. Filters such as the notch filter [17] and Butterworth filter [18] are also configured during setup. Equal gender representation is ensured by recruiting 15 participants—seven males and eight females—ranging in age from 18 to 25 years. A key inclusion criterion is short hair, as longer or coarser hair can impede electrode-skin contact, resulting in poor signal quality and increased artifacts due to a lower signal-to-noise ratio (SNR).

The preprocessing pipeline is illustrated in Fig. 4. The collected EEG data are affected by artifacts arising from muscular activity, eye blinks, and environmental noise. To enhance signal quality, preprocessing is performed following the selection of a reference channel. Given that EEG signals span a broad frequency range; a sequence of filters is employed to isolate neural activity of interest. Low-pass filters attenuate high-frequency noise originating from power lines, whereas high-pass filters eliminate baseline drift and low-frequency fluctuations.

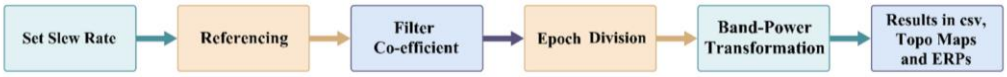



Fig. 4 Preprocessing pipeline

A simple threshold-based method is employed for additional artifact removal. To normalize the signals, a band-pass filter is applied to each windowed EEG segment. The filtered output is divided by the window length, and the Hanning window is multiplied by two to compensate for its attenuation effect. The direct current (DC) component is excluded to prevent distortion of low-frequency components in the fourier domain. Finally, the mean is subtracted from each epoch to remove offsets and ensure consistent baseline alignment across segments.

Before inserting the raw EEG data into the machine learning model, the signals undergo a transformation using CWT. This approach converts each temporal signal into a scalogram representation, as illustrated in Table 2. The primary advantage of CWT lies in its ability to provide a multi-resolution analysis of the EEG signals, capturing both time-domain and frequency-domain characteristics simultaneously.

By applying CWT, subtle features and transient oscillatory patterns within the EEG signals are effectively extracted. These rich representations enhance the model's ability to recognize intricate signal variations associated with different imagined speech commands. The resulting scalogram serve as high-dimensional input features for training and classification using machine learning algorithms. Table 2 shows an example of a scalogram image produced through CWT, reflecting the nuanced structure of the EEG activity.

### *2.5. Command training protocol*

To effectively train the system to classify four imagined commands—push, pull, left, and right—each participant engages in a structured recording protocol. The session begins with the collection of neutral EEG data, during which participants are

instructed to relax and perform controlled breathing exercises. This ensures a reliable baseline signal with minimal cognitive or muscular interference.

For each active command, participants are instructed to engage in motor imagery by vividly visualizing the sensation and muscular tension associated with performing the target action. For the “Left” and “Right” commands in particular, participants are instructed to mentally simulate the directional pulling sensation to reinforce lateral movement intent.

Each participant repeats the imagined command visualization for 8 seconds per trial, completing 10 trials per command. This repetition helps stabilize cognitive engagement and ensures consistency in signal quality. All trials are performed under standardized environmental conditions to reduce noise and artifacts. This ensures the reliability and viability of the acquired EEG data for model training and evaluation.

### *2.6. Multichannel EEG representation and time–frequency feature extraction*

Following preprocessing, the EEG signals are represented as sinusoidal waveforms, visualized in Fig. 5. To further enhance the representation and extract high-quality temporal and frequency-domain information, the CWT is applied. Unlike the traditional fourier transform, which provides only frequency information, CWT retains both frequency and time-localization. This makes it well-suited for non-stationary biomedical signals such as EEG.

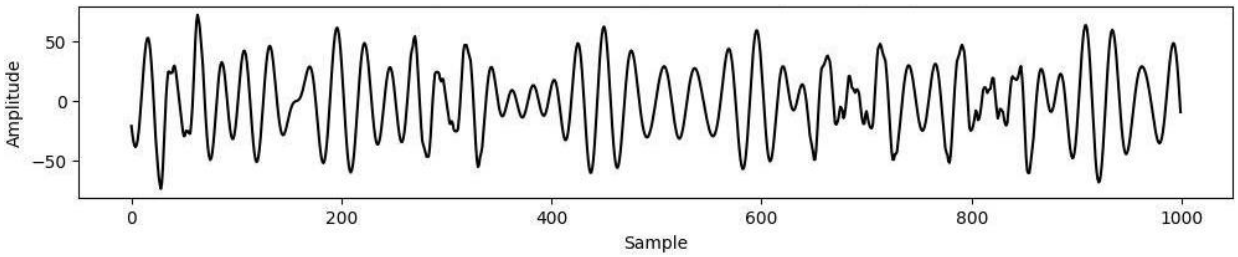



Fig. 5 Pre-processed EEG signal from the AF3 channel during the first epoch

The raw EEG signals acquired from the scalp are often noisy and incomplete representations of true brain activity. Therefore, initial filtering steps remove common sources of interference such as power line noise (50/60 Hz) using notch filters [17], and muscle artifacts using bandpass filtering (typically 0.1 – 40 Hz). For any EEG channel *n*, such as AF3, the signal at time *t* can be modeled as a combination of multiple sinusoidal components. Each component represents a different brainwave frequency relevant to the cognitive task. The EEG signal for channel AF3 can be expressed using the following formula:

$$X_{AF3}(t) = \sum_{i=1}^{n} A_i \, sin\,(2\pi f_i t + \phi_i) \tag{1}$$

where $X_{AF3}(t)$ is the EEG signal for the AF3 channel at time t, A is the amplitude of the i-th sinusoidal component, F is the frequency of the i-th sinusoidal component, $\phi$ is the phase shift of the i-th sinusoidal component, *n* is the number of sinusoidal components considered. The EEG signals from all 14 channels (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4) are structured as a multivariate time-series. At each time point *t*, the signals from all electrodes are arranged into a column vector, as shown in the following formula:

$$z(t) = [X_{AF3}(t)\, X_{AF7}(t) \ldots X_{AF4}(t)\ ] \tag{2}$$

where $x_c(t)$ represents the EEG signal from channel *c* at time *t.* The sequence of these vectors over time forms the data matrix $X(t) = \epsilon R^{14 \text{ x } \epsilon\epsilon T}$, which serves as the input for the preprocessing and feature extraction pipeline.

To extract meaningful features for classification, each channel signal $x_c(t)$ is subjected to a CWT using the complex morlet wavelet. This approach provides a time-frequency decomposition suited for non-stationary EEG signals. The CWT can be expressed by the following formula:

$$W_x(a,b) = \frac{1}{|a|^{\frac{1}{2}}} \int_{\infty}^{-\infty} x(t) \underline{\psi}\left(\frac{t-b}{a}\right) dt \quad (3)$$

where $\underline{\Psi}(t)$ is the morlet mother wavelet, a denotes the scale (inverse of frequency), and b denotes time translation. The morlet wavelet provides an optimal trade-off between time and frequency resolution, making it particularly suitable for analyzing non-stationary EEG components. The resulting wavelet coefficients $W_x$ (a.b) are then used to compute energy-based features across relevant frequency theta and alpha bands, enabling a robust representation of cognitive activity during imagined speech tasks.

Algorithm 1 EEG Signal Processing and Transformation

Input: Raw EEG signal data from 14 channels

Preprocessing: Filter out power line noise (50/60 Hz).

Apply notch filters for EMG noise removal. Apply bandpass filtering (0.1-40 Hz).

Signal Representation (AF3 Channel):

$$\chi_{AF3}(t) \leftarrow \sum_{i=1}^{n} A_i \sin(2\pi f_i t + \phi_i)$$

where: $A_i$ is amplitude of i-th component, $f_i$ is the frequency of i-th component, $\phi_i$ is the phase shift of the i-th component. $n$ is the number of sinusoidal components.

Channel Matrix Formation:

$$\chi(t) \leftarrow [\chi_{AF3}(t), \chi_{F7}(t), \ldots, \chi_{AF4}(t)]^T$$

Wavelet Transform: For each channel:

$$W_c(a,b) \leftarrow \frac{1}{|a|} \int_{-\infty}^{\infty} \chi(t) \psi^*\left(\frac{t-b}{a}\right) dt$$

where: $\psi^*$ is a complex conjugate of the mother wavelet, $a$ is a scale parameter, $b$ is a translation parameter.

Scalogram Computation:

$$S(a,b) \leftarrow |W_c(a,b)|^2$$

return Processed signal representation and scalogram.

Algorithm 2 EEG Signal Processing and Scalogram Generation

Input: Raw EEG signals $X_{raw}$ from 14 channels

Output: Scalogram S(a,b)

Procedure PREPROCESS($X_{raw}$)

Apply notch filter $F_{notch}(50/60\ Hz)$ to remove powerline noise

Apply bandpass filter $F_{band}(0.1/40\ Hz)$ to suppress EMG artifacts

For each channel $c \in \{AF3, F7, F3, \ldots, AF4\}$ do

Decompose signal into sinusoidal components:

$$x_c(t) = \sum_{i=1}^{n} A_i sin(2\pi f_i t + \phi_i)$$

End for

Construct multichannel EEG vector:

$$z(t) = [X_{AF3}(t)\ X_{AF7}(t) \ldots X_{AF4}(t)\ ]$$

### *2.7. System deployment and communication architecture*

For the practical implementation of the proposed BCI system, the Emotiv Epoc X 14-channel EEG headset is chosen for its wireless transmission capabilities and reliable signal acquisition. A notable communication challenge arises from the proprietary architecture of the Cortex API, which does not support direct signal transmission to embedded hardware such as the Raspberry Pi.

To overcome this limitation, a laptop is configured as a communication bridge between the EEG headset and the Raspberry Pi onboard the robot. Both the laptop and Raspberry Pi are connected to the same local area network (LAN), enabling seamless real-time communication. In this setup, EEG data are collected, pre-processed, and classified on the laptop. The resulting command is then transmitted to the Raspberry Pi with a latency of less than 0.5 milliseconds using the WebSocket protocol.

The Cortex Python SDK is employed on the server side to establish and authenticate a connection with the headset. Authentication requires a valid license key obtained through the official Emotiv developer portal. The WebSocket protocol is selected for its low overhead and persistent connection properties, ensuring reliable, continuous data transmission to the robotic controller. This architecture as shown in Fig. 6, enables real-time operation and accurate translation of neural commands into robotic movements.

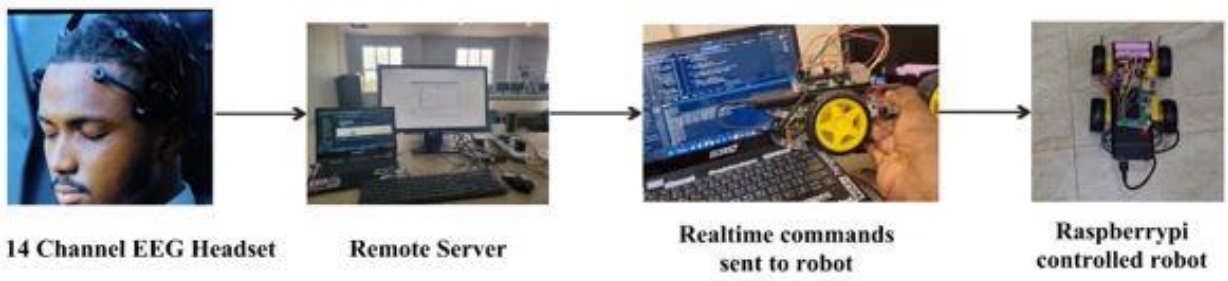



Fig. 6 Implemented BCI framework

### *2.8. Robot platform design and control implementation*

To demonstrate the practical functionality of the proposed BCI system, a robotic platform is constructed and controlled via a Raspberry Pi microcontroller. The Raspberry Pi 4 is chosen for its balance of computational performance, compact form factor, and extensive general-purpose input/output (GPIO) pin support. With 26 GPIO pins available, the board offers ample connectivity for motor drivers and peripheral sensors, making it an ideal core processor for the robotic control system.

Physical integration focuses on interfacing the Raspberry Pi with the robot's motion subsystems. An L298N H-bridge motor driver serves as the intermediary between the Raspberry Pi's low-voltage logic signals and the higher-current demands of the DC motors. This driver facilitates bidirectional motor control, enabling precise movement in forward, reverse, left, and right directions based on neural input.

The robot is powered by a 7.4 V battery, supplying sufficient energy for extended wireless operation. GPIO pins on the Raspberry Pi are connected to the L298N driver according to a predefined wiring configuration, ensuring consistent signal routing for all motor control operations.

On the software side, a Python-based control script translates EEG classification results into actionable motor commands. The program initialized GPIO configurations and implements a real-time control loop that continuously listened for incoming data via a WebSocket connection. Each recognized neural command—push, pull, left, right, or neutral—is mapped to a corresponding motor action, enabling intuitive and responsive control of the robot in real time.

## 3. Results

This section presents the comprehensive results of the BCI system evaluation, including classifier performance, model comparison against existing literature, and detailed time-frequency analysis of the imagined commands. The primary objective is to validate the system's capability to accurately decode four imagined speech commands (push, pull, left, right) and a neutral state in real-time.

### *3.1. Classifier performance evaluation*

System performance is evaluated along four dimensions: classification accuracy, cross-subject robustness, and comparative model performance. Classification accuracy is measured as the proportion of correctly decoded imagined speech commands. Cross-subject performance is assessed by examining model accuracy across different recording sessions and participants.

To identify the most effective algorithm for EEG signal classification, a range of supervised learning models is evaluated, including decision trees, logistic regression, SVM, KNN, random forests, gradient boosting machines and XGBoost. These models are compared based on their classification accuracy across the four imagined commands, using features extracted from recordings of five randomly selected subjects.

The dataset is split using an 80/20 train-test split with a random state of 72 for reproducibility. This yields a training set of 164,412 samples (80%) and a test set of 41,103 samples (20%). No feature scaling is applied, as XGBoost is scale-invariant. XGBoost is configured with a learning rate of 0.05 and a Synthetic Minority Over-sampling Technique (SMOTE) oversampling on the training set to address class imbalance *scale_pos_weight* calculated as the ratio of majority to minority class frequencies. XGBoost achieves an accuracy of 98%, a weighted precision of 98.51%, and a weighted F1-Score of 98.43%.

The neutral state is used for baseline calibration and signal stabilization rather than as an explicit command class. The results, displayed in Fig. 7 and summarized in Table 3, demonstrate that XGBoost outperforms all other classifiers, achieving an average accuracy of approximately 98% across the five command classes, including the neutral state. This outcome highlights the strength of the feature extraction pipeline and the model’s robustness in managing EEG signal variability associated with imagined speech classification.

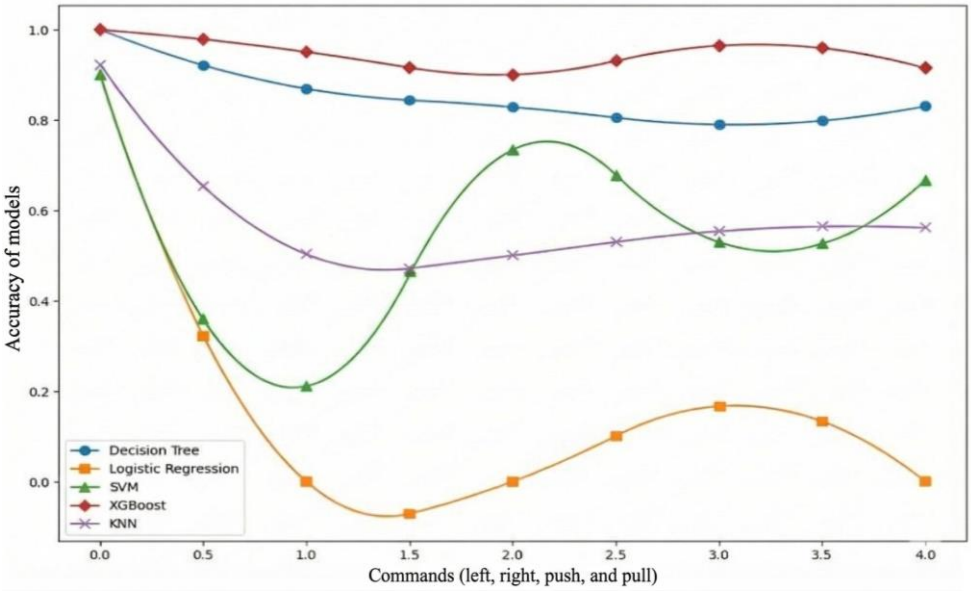



Fig. 7 Comparison of XGboost against other models

Table 3 Accuracy of commands by subject and model sample

| Push Command | | | | | |
|---|---|---|---|---|---|
| Command | Subject t-1 | Subject t-2 | Subject t-3 | Subject t-4 | Subject t-5 |
| Decision Tree | 0.87 | 0.33 | 0.89 | 0.95 | 0.97 |
| Logistic Regression | 0.00 | 0.25 | 0.00 | 0.00 | 0.00 |
| Random Forests | 1.00 | 1.00 | 1.00 | 0.93 | 0.99 |
| SVM | 0.21 | 0.99 | 0.60 | 0.65 | 0.99 |
| KNN | 0.50 | 0.35 | 0.71 | 0.71 | 0.61 |
| XGBoost (Proposed) | 0.95 | 0.0 | 0.95 | 0.93 | 0.98 |
| Pull Command | | | | | |
| Command | Subject t-1 | Subject t-2 | Subject t-3 | Subject t-4 | Subject t-5 |
| Decision Tree | 0.83 | 0.81 | 0.97 | 0.84 | 0.84 |
| Logistic Regression | 0.00 | 0.00 | 0.35 | 0.00 | 0.58 |
| Random Forests | 1.00 | 1.00 | 1.00 | 0.99 | 0.93 |
| SVM | 0.67 | 0.50 | 0.77 | 0.61 | 0.74 |
| KNN | 0.56 | 0.67 | 0.72 | 0.63 | 0.74 |
| XGBoost (Proposed) | 0.91 | 0.97 | 0.99 | 0.92 | 0.92 |
| Left Command | | | | | |
| Command | Subject t-1 | Subject t-2 | Subject t-3 | Subject t-4 | Subject t-5 |
| Decision Tree | 0.83 | 0.93 | 0.75 | 0.64 | 0.75 |
| Logistic Regression | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Random Forests | 1.00 | 1.00 | 1.00 | 10.0 | 1.00 |
| SVM | 0.73 | 0.33 | 0.22 | 0.00 | 0.23 |
| KNN | 0.50 | 0.35 | 0.00 | 0.43 | 0.67 |
| XGBoost (Proposed) | 0.90 | 0.98 | 0.100 | 0.100 | 0.94 |
| Right Command | | | | | |
| Command | Subject t-1 | Subject t-2 | Subject t-3 | Subject t-4 | Subject t-5 |
| Decision Tree | 0.79 | 0.94 | 1.00 | 0.85 | 0.75 |
| Logistic Regression | 0.17 | 0.00 | 0.00 | 0.00 | 0.00 |
| Random Forests | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| SVM | 0.52 | 0.23 | 1.00 | 0.45 | 0.67 |
| KNN | 0.55 | 0.67 | 0.33 | 0.75 | 1.00 |
| XGBoost (Proposed) | 0.96 | 0.94 | 0.83 | 1.00 | 1.00 |
| Neutral Command | | | | | |
| Command | Subject t-1 | Subject t-2 | Subject t-3 | Subject t-4 | Subject t-5 |
| Decision Tree | 1.00 | 1.00 | 0.92 | 0.94 | 0.73 |
| Logistic Regression | 0.90 | 0.99 | 0.92 | 0.94 | 0.73 |
| Random Forests | 1.00 | 0.99 | 0.99 | 0.99 | 1.00 |
| SVM | 0.90 | 0.99 | 0.94 | 0.95 | 0.84 |
| KNN | 0.92 | 0.99 | 0.96 | 0.98 | 0.83 |
| XGBoost (Proposed) | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |


Fig. 8 Confusion matrix of XGboost model

To further assess the model's reliability, the XGBoost classifier's confusion matrix is presented in Fig. 8. The results show that the model successfully differentiated among the five commands with minimal errors. Some misclassification occurs for the "Pull" command, with 130 instances erroneously labelled as "Push." This indicates that targeted data augmentation or algorithmic calibration could improve class-specific precision.

Each EEG session per subject is limited to 10 minutes. Consequently, some neural patterns may not be fully captured within that timeframe. Longer and more frequent recording sessions could provide more comprehensive data and support better generalization in future studies.

For the neutral class (label 0), the model achieves perfect accuracy with 3,800 correct classifications and no misclassifications. This indicates excellent sensitivity and specificity in detecting resting-state brain activity.

The "Push" command (label 1 horizontally) is correctly classified in 3,900 cases. However, it is occasionally confused with the "Right" command (130 instances), and less frequently with "Pull" (23 instances) and "Left" (7 instances), indicating potential overlap in the neural activation patterns associated with forward and backward motor imagery.

For the "Left" command (label 2 horizontally), the model again achieved 3,900 correct predictions but exhibited confusion with "Right" (110 instances), "Push" (48 instances), and "Left" (11 instances), reflecting shared cognitive features across lateral and axial imagery.

The "Right" command (label 3 horizontally) yielded 2,000 correct classifications, with misclassifications into "Right" (120 instances) and "Push" (19 instances), suggesting the neural signature for rightward intention may partially overlap with other directions.

The "Pull" command (label 4) showed the fewest correct classifications (1,200) and the highest confusion rate, especially with "Push" and "Left," indicating that additional data or enhanced feature extraction techniques may be needed to improve model sensitivity for this class. These findings underscore the effectiveness of the model while highlighting areas for improvement through expanded training data, class balancing, and task-specific calibration.

### *3.2. Performance Comparison with EEG-Based Studies*

To contextualize the performance of the proposed BCI system, it is compared against several recent EEG-based classification frameworks developed between 2022 and 2024 [19-25]. These studies represent the current state-of-the-art in imagined speech and motor imagery decoding using EEG signals. Table 4 presents a structured comparison based on classification accuracy, EEG channel count, classifier used, and whether real-time deployment is demonstrated.

Table 4 Comparison of EEG classification methods

| Study | Year | Task Type | Dataset | Subjects | Channels | Feature Extraction | Classifier | Accuracy (%) | Real Time |
|---|---|---|---|---|---|---|---|---|---|
| [19] | 2023 | M+I | – | 9 | 22 | T-D+ | TCN | 88.40 | False |
| [20] | 2023 | I+S | C | 10 | 32 | dWT | CNN | 84.50 | False |
| [21] | 2024 | V | SUD | 10 | — | LSTM | SVM | 66.67 | False |
| [22] | 2024 | M+I | BCI-IV-2a | 9 | 22 | TD-LSTM | LSTM | 99.09 | False |
| [23] | 2021 | M+I | — | — | — | MB | CNN | 99.85 | False |
| [24] | 2024 | D-D | — | — | — | BM | AI-based | 99.75 | False |
| [25] | 2025 | I+S | CI | 20 | 64 | S+A | Transformer | 91.80 | False |
| This Work | 2025 | I+S | C | 15 | 14 | CWT+S | XGBoost | 98.10 | True |

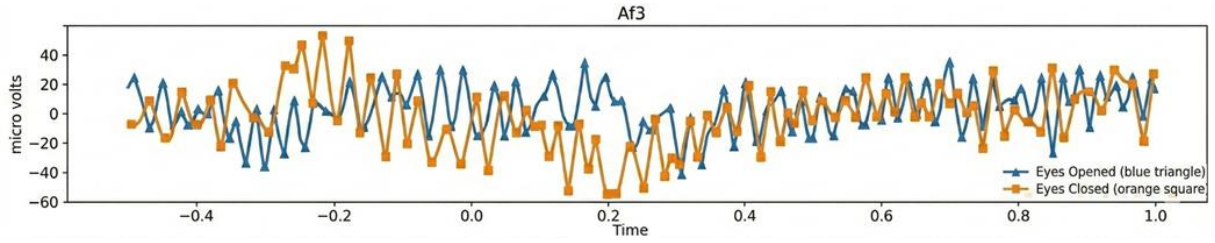
**Note:** *a.* I+S: *Imagined Speech, b.* M+I: *Motor Imagery, c.* V: *Visual task, d.* CWT: *Continuous Wavelet Transform and e.* S+A: *Spectrogram with Attention*

Unlike prior studies that focus on offline imagined speech decoding or rely on high-density clinical EEG systems, this work integrates imagined speech decoding with robotic actuation in a single real-time BCI framework. The system uses a consumer-grade EEG headset and embedded hardware. The proposed approach achieves high classification accuracy while using fewer EEG channels and supporting real-time deployment. These results emphasize the system's efficiency and highlight its suitability for low-cost, practical applications in assistive robotics.

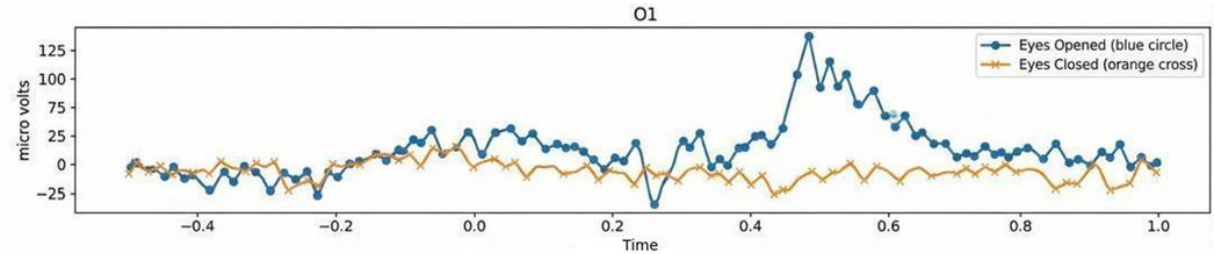
### *3.3. Event-related potentials* (*ERPs*) *analysis*

ERPs are characterized by sequences of positive and negative voltage deflections occurring within specific time windows following a stimulus, as described by McWeeny and Norton [26]. These components provide a millisecond-by-millisecond view of cognitive processes, offering insight into how the brain perceives, evaluates, and responds to external inputs.

In this study, ERP analysis is employed to determine the temporal order in which different cognitive processes are activated in response to motor imagery commands. By examining the timing, amplitude, and scalp distribution of ERP peaks, the brain's sensory and cognitive response mechanisms are assessed.




(a) A graphical representation of the AF3 (frontal) channel




(b**)** A graphical representation of the O1 (occipital) channel

Fig. 9 The Time- and frequency-domain analysis of EEG signals under Eyes Opened and Eyes Closed conditions

Early ERP components (within the first 100 milliseconds post-stimulus) typically reflect initial sensory encoding and registration of stimuli within their respective sensory modalities (e.g., visual or auditory). Later components (between 100 and 500 milliseconds) are associated with higher-level processes, such as attention allocation, feature discrimination, categorization, and decision-making. Analyzing these later-stage ERPs allows us to assess participants' attention to the task and their preparation to issue the correct mental command. Fig. 9 illustrates the average grouped ERP waveforms recorded from a representative subject across multiple trials.

Time-domain EEG signals from the AF3 (frontal) and O1 (occipital) channels show marked differences between the Eyes Opened and Eyes Closed conditions. Frontal signals exhibit increased theta and alpha modulation during visual engagement. In contrast, occipital signals show elevated alpha activity during visual disengagement (eyes closed), confirming region-specific spectral dynamics.

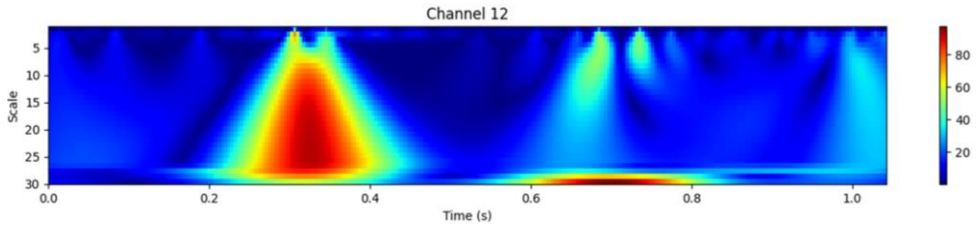
In the AF3 channel, the Eyes Opened condition presents greater amplitude fluctuations, indicating increased frontal lobe activity associated with visual attention and cognitive engagement. In the O1 channel, a significant amplitude peak during the Eyes Opened state suggests strong occipital activation due to visual stimulus processing.

To further explore these dynamics, continuous wavelet transform (CWT) using the morlet wavelet is applied, revealing frequency-specific energy distributions over time. The frontal AF3 region demonstrated increased power in the theta (4–8 Hz) and lower alpha (8–10 Hz) bands during the Eyes Opened state. Meanwhile, the O1 region showed enhanced alpha activity (8–12 Hz) when the eyes are closed—a well-known marker of visual cortical idling. These findings validate the time-frequency sensitivity of the morlet transform and highlight the importance of both temporal and spectral features in developing robust EEG-based BCI systems for imagined speech or attention-state decoding.

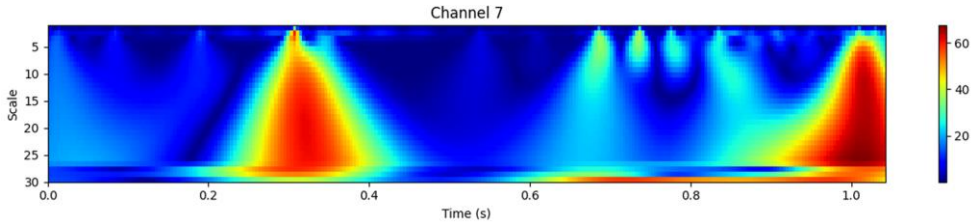
### *3.4. Time–frequency analysis of imagined commands using the morlet wavelet*

(1) Imagined “Push” Command

Time-frequency representations derived from morlet wavelet transforms for Channels 7 and 12 are illustrated in Fig. 10. Channel 12 exhibits a sharp and intense activation in the lower frequency range (high scale values) around 0.35–0.45 seconds, indicating a focused neural response likely tied to cognitive engagement during the imagined “push” task. In contrast, Channel 7 demonstrates a broader and more dispersed activation pattern with power distributed over multiple time windows and frequency bands. This distinction suggests regional variation in neural processing strategies involved in motor imagery.




(a) Morlet wavelet time–frequency representation of Channel 12

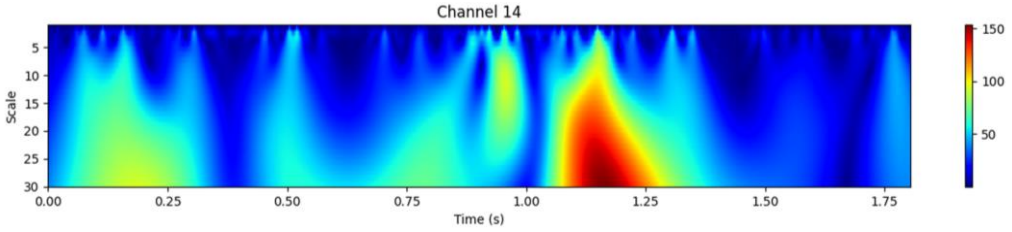



(b) Morlet wavelet time–frequency representation of Channel 7

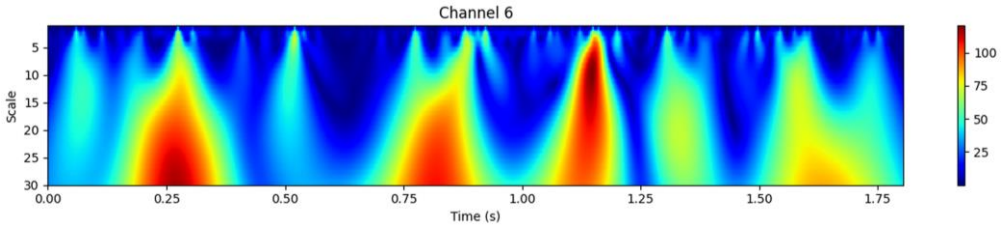
Fig. 10 Time-frequency scalograms of Channel 7 and Channel 12 during the "Push" command

(2) Imagined “Pull” Command

Time-frequency plots for Channels 6 and 14 during the imagined “pull” command are shown in Fig. 11. Channel 14 exhibits a significant and sharply localized activation in lower frequencies between 1.0 and 1.3 seconds, implying strong regional involvement in the motor imagery process. Conversely, Channel 6 demonstrates periodic bursts of activity across a broader temporal window and frequency range, indicating more distributed neural dynamics. These findings reinforce the hypothesis that different cortical regions contribute distinctively to specific motor imagery tasks.




(a) Morlet wavelet time–frequency representation of Channel 14

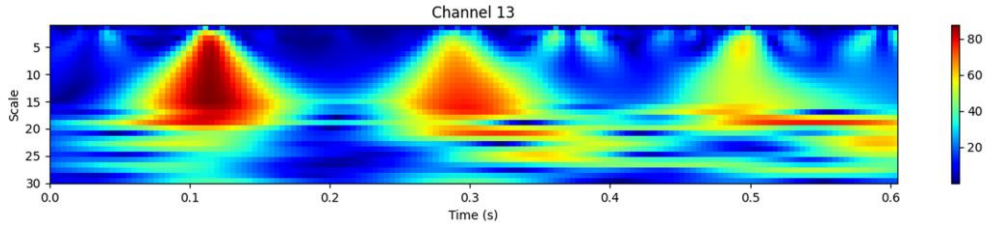



(b) Morlet wavelet time–frequency representation of Channel 6

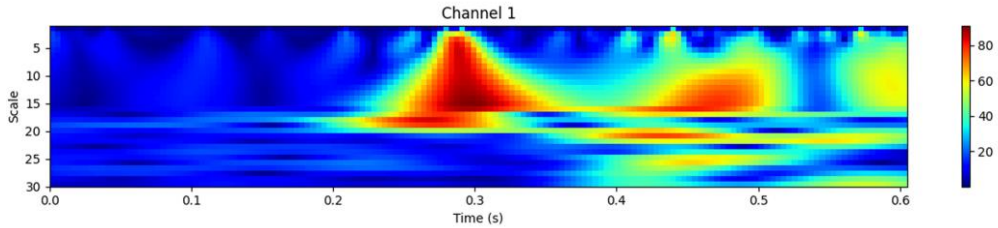
Fig. 11 Time-frequency scalograms of Channel 6 and Channel 14 during the "Pull" command

(3) Imagined “Left” Command

Time–frequency representations for Channels 13 and 1 during the imagined “left” command are presented in Fig. 12. Channel 13 displayed two distinct and temporally early low-frequency activations around 0.1 and 0.3 seconds, suggesting rapid onset of cognitive processes involved in directional motor imagery. These peaks are prominent in lower frequency bands, indicating engagement of cortical areas associated with preparatory and planning phases of movement. In contrast, Channel 1 demonstrated a more concentrated activation centered at approximately 0.3–0.4s with extended spectral spread toward higher scales, suggesting sustained engagement of motor-related cortical regions. Together, these patterns highlight temporal and spatial variability in neural activation during imagined leftward movements.




(a) Morlet wavelet time–frequency representation of Channel 13

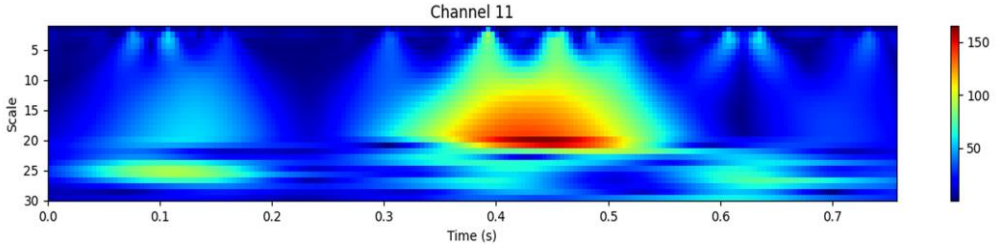



(b) Morlet wavelet time–frequency representation of Channel 1

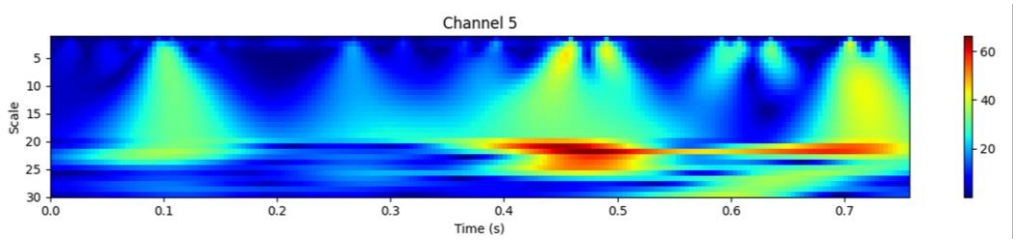
Fig. 12 Time-frequency scalograms of Channel 1 and Channel 13 during the "left" command

(4) Imagined "Right" Command

Time–frequency analyses for Channels 11 and 5 during the imagined "right" command are illustrated in Fig. 13. Channel 11 exhibits a distinct, concentrated activation in the lower frequency range (higher scales) around 0.4–0.5 seconds, indicating a focal neural response likely linked to motor planning or execution components of the imagined action. In comparison, channel 5 demonstrates a broader activation in the same frequency range beginning slightly earlier (approximately 0.35s) and extending toward 0.55s. The distributed temporal spread of Channel 5's activation suggests involvement in complementary preparatory or integrative neural processes. Collectively, these observations support the notion of coordinated yet regionally diverse neural engagement during imagined rightward movement.




(a) Morlet wavelet time–frequency representation of Channel 11




(b) Morlet wavelet time–frequency representation of Channel 5

Fig. 13 Time-frequency scalograms of Channel 11 and Channel 5 during the "right" command

### *3.5. Limitations*

While the proposed BCI system achieved high accuracy and demonstrated effective real-time robotic control, several limitations must be acknowledged. First, the dataset is limited to 15 participants, which may not fully capture inter-subject variability in EEG signal patterns. This constraint could affect the system's generalizability to broader populations.

Second, the command vocabulary is restricted to four imagined speech actions (push, pull, left, right), limiting the system's utility for more complex tasks. Expanding the command set and incorporating continuous speech or multi-lingual imagined phrases would significantly enhance usability.

Although the system supports real-time control using a Raspberry Pi, performance under varied hardware conditions, environmental noise, and user stress levels remains unexplored. The impact of cognitive load and fatigue on system accuracy also warrants further investigation.

## 4. Conclusions

This study demonstrates the feasibility and effectiveness of a non-invasive BCI system capable of translating EEG signals into real-time robotic control with high accuracy. The integration of advanced preprocessing techniques, time–frequency feature extraction using CWT, and the XGBoost classifier enabled the system to achieve a classification accuracy of 98%.

While the current system supports a limited set of commands—push, pull, left, and right—its design provides a foundation for future scalability. Further research can expand the command set, explore more complex imagined speech patterns, and enhance model generalizability through larger and more diverse datasets.

This innovation has significant potential in assistive and rehabilitative applications, particularly for individuals with severe motor impairments. By enabling hands-free, thought-driven control of robotic platforms, the system opens pathways to improved independence and quality of life.

As BCI research continues to evolve, this work contributes toward the broader goal of developing intuitive, accessible, and practical BCI systems for real-world use. Future efforts should also address domain adaptation, user personalization, and integration with multimodal feedback systems to further advance the capabilities of BCI-driven technologies.

Future work will focus on:

(1) Expanding the dataset to include more diverse subjects for better generalizability.
(2) Integrating emotion detection and adaptive feedback mechanisms [26].
(3) Exploring deep learning and hybrid models to improve classification as described by Ahmadieh et al. [27].
(4) Testing the system across different real-world assistive scenarios [28].

## Conflicts of Interest

The authors declare no conflict of interest.

## References

[1] S. R and S. S Tippannavar, "EEG Based Smart Wheelchair Using Raspberry Pi for Elderly and Paralysed Patients," Proceedings of the 2nd IEEE Mysore Sub Section International Conference (MysuruCon), IEEE press, pp. 1-5, 2022.
[2] V. Chaurasia, V. Mishra, and L. Jain, "Brain-Bot: An Unmanned Ground Vehicle (UGV) Using Raspberry Pi and Brain Computer Interface (BCI) Technology," Proceedings of Advances in Computing and Data Sciences: ICACDS 2016, Communications in Computer and Information Science, vol. 721, pp. 252-261, 2017
[3] I. Rakhmatulin, and S. Volkl, "Brain-Computer-Interface Controlled Robot via RaspberryPi and PiEEG," arXiv preprint arXiv:2202.01936, 2022.
[4] V. Asanza, A.Constantine, S. Valarezo, and E. Peláez, "Implementation of a Classification System of EEG Signals Based on FPGA," Proceedings of Seventh International Conference on eDemocracy & eGovernment (ICEDEG), pp. 87-92, 2020.
[5] B. B. Das, P. Kumar, D. Kar, S. K. Ram, K. S. Babu, and R.K.Mohapatra, "A Spatio-Temporal Model for EEG-Based Person Identification," Multimedia Tools and Applications, vol. 78, no. 19, pp. 28157-28177, 2019.
[6] E. H. Houssein, A. Hammad, and A. A. Ali, "Human Emotion Recognition from EEG-Based Brain –Computer Interface Using Machine Learning: A Comprehensive Review," Neural Computing and Applications, vol. 34, no. 15, pp. 12527-12557, 2022.
[7] P. Saha, S. Fels, and M. Abdul-Mageed, "Deep Learning the EEG Manifold for Phonological Categorization from Active Thoughts," Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2762-2766, 2019.
[8] P. Kumar, and E. Scheme, "A Deep Spatio-Temporal Model for EEG-Based Imagined Speech Recognition," Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 995-999, 2021.
[9] A. Chaudhry, U. Khan, M. R. Palla, S. B. Singh, and S. V. Deshmukh, "A Prosthetic Arm Based on Electroencephalography by Signal Acquisition and Processing on MATLAB," International Journal of Research in Engineering, Science and Management, vol. 5, no. 1, pp. 119-124, 2022.
[10] W. Zgallai, J. T. Brown, A. Ibrahim, F. Mahmood, K. Mohammad, and Maitha Khalfan, "Deep Learning AI Application to an EEG Driven BCI Smart Wheelchair." Proceedings of Advances in Science and Engineering Technology International Conferences (ASET), pp. 1-5, 2019.
[11] R. Bousseta, I. El Ouakouak, M. Gharbi, and F. Regragui, "EEG Based Brain Computer Interface for Controlling a Robot Arm Movement Through Thought," IRBM, vol. 39, no. 2, pp. 129-135, 2018.
[12] M. AlSaleh, R. Moore, H. Christensen, and M. Arvaneh, "Discriminating Between Imagined Speech and Non-Speech Tasks Using EEG," Proceedings of 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1952-1955, 2018.
[13] U. Lal, A. V. Chikkankod, and L. Longo, "Fractal Dimensions and Machine Learning for Detection of Parkinson's Disease in Resting-State Electroencephalography," Neural Computing and Applications, vol. 36, no. 15, pp. 8257-8280, 2024.

[14] Y. An, and H. K. Lam, and S. H. Ling, "Multi-Classification for EEG Motor Imagery Signals Using Data Evaluation-Based Auto-Selected Regularized FBCSP and Convolutional Neural Network," Neural Computing and Applications, vol. 35, no. 16, pp. 12001-12027, 2023.

[15] D. Raab, A.Theissler, and M. Spiliopoulou, "XAI4EEG: Spectral and Spatio-Temporal Explanation of Deep Learning-Based Seizure Detection in EEG Time Series," Neural Computing and Applications, vol. 35, no. 14, pp. 10051-10068, 2023.

[16] X. Xu, N. Drougard, and R. N. Roy, "Topological Data Analysis as a New Tool for EEG Processing," Frontiers in Neuroscience, vol. 15, article no. 761703, 2021.

[17] K. Hirano, S. Nishimura, and S. Mitra, "Design of Digital Notch Filters," IEEE Transactions on Communications, vol. 22, no. 7, pp. 964-970, 1974.

[18] I. W. Selesnick, and C. S. Burrus, "Generalized Digital Butterworth Filter Design," IEEE Transactions on Signal Processing, vol. 46, no. 6, pp. 1688-1694, 1998.

[19] I. Aliyu, and C. G. Lim, "Selection of Optimal Wavelet Features for Epileptic EEG Signal Classification with LSTM," Neural Computing and Applications, vol. 35, no. 2, pp. 1077-1097, 2023.

[20] R. M. Rady, D. Elsalamawy, M. R. M. Rizk, O. A. Alim, and N. D. Moussa, "Enhancing Affordable EEG to Act as a Quantitative EEG for Inattention Treatment Using MATLAB," Neural Computing and Applications, vol. 37, pp. 5849-5871, 2025.

[21] M. Karimian-Kelishadrokhi, and F. Safi-Esfahani, "TD-LSTM: A Time Distributed and Deep-Learning-Based Architecture for Classification of Motor Imagery and Execution in EEG Signals," Neural Computing and Applications, vol. 36, no. 25, pp. 15843-15868, 2024.

[22] G. Kumar, T. Das, and K. Singh, "Early Detection of Depression through Facial Expression Recognition and Electroencephalogram-Based Artificial Intelligence-Assisted Graphical User Interface," Neural Computing and Applications, vol. 36, no. 12, pp. 6937-6954, 2024.

[23] M. Nour, S. Ozturk, and K. Polat, "A Novel Classification Framework Using Multiple Bandwidth Method with Optimized CNN for Brain-Computer Interfaces with EEG-fNIRS Signals," Neural Computing and Applications, vol. 33, no. 22, pp. 15815-15829, 2021.

[24] R. Saia, S. Carta, G. Fenu, and L. Pompianu, "Influencing Brain Waves by Evoked Potentials as Biometric Approach: Taking Stock of the Last Six Years of Research," Neural Computing and Applications, vol. 35, no. 16, pp. 11625-11651, 2023.

[25] T. Khanam, S. Siuly, and H. Wang, "An Optimized Artificial Intelligence Based Technique for Identifying Motor Imagery from EEGs for Advanced Brain Computer Interface Technology," Neural Computing and Applications, vol. 35, no. 9, pp. 6623-6634, 2023.

[26] S. McWeeny, and E. Norton, "Understanding Event-Related Potentials (ERPs) in Clinical and Basic Language and Communication Disorders Research: A Tutorial," International Journal of Language & Communication Disorders, vol. 55, no. 4, pp. 445-457, 2020.

[27] H. Ahmadieh, F. Gassemi, and M. H. Moradi, "A Hybrid Deep Learning Framework for Automated Visual Image Classification Using EEG Signals," Neural Computing and Applications, vol. 35, no. 28, pp. 20989-21005, 2023.

[28] X. Y. Liu, W. L. Wang, M. Liu, and M. Y. Cheni, T. Pereira, D. Y. Doda, Y. F. Ke, et al., "Recent Applications of EEG-Based Brain-Computer-Interface in the Medical Field," Military Medical Research, vol. 12, no. 14, pp. 1-42, 2025.