

Human–IoT Interaction via EEG-Based Brain–Computer Interface Using Weak AI Models

Mohamed A. Gismelbari, Ilya I. Vixnin, A. Alhasan, Gataullin R.I.

Saint Petersburg Electrotechnical University “LETI”,

Saint Petersburg, Russian Federation

mohamedgbari@gmail.com (M.A.G), wixnin@mail.ru (I.I.V), Aliyossefahasan@gmail.com (A.A), grintez@gmail.com (G.R.I).

Abstract— In this paper, we explore and validate the feasibility of using electroencephalography (EEG) based brain–computer interfaces (BCIs) to issue basic control commands to unmanned aerial vehicles (UAVs). We focus on integrating human cognitive motor commands with Internet of Things (IoT) devices, enabling hands-free UAV control. In our approach, neural signals captured during motor imagery of a right-hand upward movement and a left-hand downward movement are translated into discrete UAV instructions (conceptually analogous to “hover” and “land” commands). EEG data were acquired from a 14-channel Emotiv Epoc X headset worn by 10 participants, and features such as band power in key frequency bands were extracted. A lightweight decision tree classifier was trained and evaluated in a leave-one-participant-out (LOPO) cross-validation scheme to assess how well the model generalizes across individuals. The results indicate that certain participants can achieve classification accuracies above 65% for the two mental commands, although average accuracy across all subjects was modest (~55%). These findings highlight both the promise and the challenges of EEG-based hands-free drone control. They demonstrate the potential of neural interfaces as a bridge between human thought and machine action in IoT contexts, while also underscoring the need for improved signal processing and personalization to handle inter-subject variability. This work lays important groundwork for more advanced BCI-driven UAV control frameworks, aiming toward intuitive human–IoT interactions in high-impact domains.

Keywords— Brain–Computer Interface (BCI); EEG; UAV Control; Motor Imagery; Neural Interface; Machine Learning; IoT; Hands-Free Operation; Emotiv Epoc.

I. INTRODUCTION

Advances in brain–computer interfaces (BCIs) and AI have opened new possibilities for direct communication between human thought and external devices in various applications, from assistive robotics to smart environments [2]. In particular, non-invasive EEG-based BCIs offer a means to integrate with IoT systems for intuitive, hands-free control of connected devices [1]. One promising direction is the use of BCI technology to control UAVs [3] via thought commands, which could be invaluable when manual control is infeasible (for instance, for users with motor impairments or in hands-busy operational scenarios) [2], [3]. Prior studies have demonstrated that both P300 event-related potential and motor imagery (MI) paradigms can be harnessed for UAV command and navigation [10], [11], [12], with reported accuracies as high as ~90% in controlled settings [4], [5]. These works show that BCIs can complement (though not yet fully replace) manual controllers by providing an alternate channel for issuing discrete commands to drones. Despite these advances, current EEG-based BCIs face challenges including relatively low bitrates, longer command

selection times, and high variability across users [6], [7]. Motor imagery-based BCIs in particular typically support only a small number of distinct commands reliably (often 2–3 classes) due to the subtlety of EEG patterns and user training requirements [8], [9]. The novelty of our work lies in introducing a streamlined approach that focuses on binary motor-imagery commands to control a UAV, specifically mapping imagined upward versus downward hand movements to basic drone instructions (“hover” and “land”). By concentrating on two fundamental commands, we aim to maximize classification reliability and minimize latency, establishing a foundational BCI control layer that could later be expanded to more degrees of freedom. Unlike the most of prior BCI drone studies that use elaborate signal processing or visual stimuli (e.g., steady-state evoked potentials), our approach leverages time-domain EEG features and a simple ML model (a decision tree). This minimalist strategy tests the baseline feasibility of a consumer-grade BCI system for drone control in real-world conditions.

In this paper, we present the methodology and results of our BCI-UAV integration. We recorded EEG signals from 10 participants performing motor imagery of two opposite hand motions and extracted a feature set comprising band-power and statistical measures from multiple EEG channels. We evaluated classification performance using a rigorous leave-one-participant-out validation to assess how well the model can generalize to unseen individuals. The results demonstrate the potential for achieving above-chance, hands-free command of a UAV through EEG signals. We discuss how our findings compare to prior work and analyze the substantial inter-subject variability observed, attributing it to factors such as individual differences in neural signatures (a phenomenon related to “BCI illiteracy” where a notable subset of users struggles with MI-BCI control) [10], [11]. The implications of limiting the system to binary commands are also examined, arguing that even a two-command BCI can be practically useful for high-level UAV supervision. Finally, we outline the limitations of the current study – including the modest accuracy and small sample size – and suggest directions for future improvements, such as incorporating adaptive algorithms or hybrid BCI paradigms to enhance reliability and command complexity.

II. MATERIALS AND METHODS

A. Participants

Ten volunteers (8 males and 2 females, ages 20–30) were recruited for the study (referred to as P01, P02, etc). All participants were right-handed and had normal or corrected-to-normal vision. None reported any neurological disorders. We

recorded basic demographics including age, gender, handedness and prior BCI experience as shown in Table I. Notably, three participants had some prior exposure to BCI use, while the others were naive first-time users. This information was collected to examine whether factors like BCI experience might influence performance. All participants gave informed consent before the experiment and were instructed on the procedure. The study focused on within-subject binary classification of imagined movements, but data from all participants were later combined to evaluate generalization across individuals. Fig.1 shows the Emotiv Epoc X setup.



Fig. 1. Device setup (Emotiv Epoc X). Placement of electrodes on the participant's head (left), downward movement (middle), upward movement (right)

TABLE I. PARTICIPANTS DEMOGRAPHICS

No.	Sex	Age	Handedness	BCI Exp.
P01	M	29	Right	No
P02	F	23	Right	No
P03	M	20	Right	Yes
P04	M	20	Right	No
P05	M	25	Right	Yes
P06	M	27	Right	No
P07	M	21	Right	No
P08	M	28	Right	Yes
P09	F	30	Right	No
P10	M	24	Right	No

B. Experimental Procedure

Participants first completed training sessions in EmotivBCI to associate imagined upward and downward movements with classifier profiles [17], [18]. The training interface provided visual feedback on success/failure within 8-second trials. An example of the EmotivBCI training interface is shown in Fig. 2.



Fig. 2. Training interface (EmotivBCI). The reference box with 8-second timer per training to execute the command “upward”, “downward”, “neutral”. Training results: goal not reached (left), goal reached (right)

Each participant engaged in a series of sessions involving MI tasks corresponding to two distinct drone control commands. The chosen mental tasks were imagining a right-hand upward motion (as if signaling a drone to rise or hover in place) and imagining a left-hand downward motion (as if signaling the drone to descend or land). These two actions were selected for their intuitive mapping to basic UAV behaviors and their

mutually exclusive nature. During an initial calibration phase, participants underwent training using the EmotivBCI software tool. This training presented visual cues and timers guiding the participant to perform or imagine the specified hand movement for a few seconds at a time, interleaved with rest (neutral) periods. Participants were instructed to vividly imagine the kinesthetic sensation of raising or lowering their arm (some participants also subtly performed the motion to reinforce the imagery) while the BCI system learned their EEG patterns for each command [10], [11]. Each mental command (“upward”, “downward”) was trained in multiple trials (on the order of 10–20 repetitions per command) until the software indicated a stable detection profile for that user [12], [13]. A neutral state (no movement/relaxed) was also recorded to serve as a baseline in the training process [7].

After calibration, each participant proceeded to the execution phase using the EmotivPro software. In this phase, live EEG data were recorded as participants attempted to issue binary commands to a virtual drone purely by performing the learned motor imagery tasks. Participants sat comfortably and, at prompted intervals, imagined the designated upward or downward hand movement without any overt physical movement. No visual flicker or P300 stimuli were used; only the participant’s self-paced motor imagery triggered the commands [4], [6]. Each session consisted of a sequence of trials where the participant was cued (visually or by an audio cue) to perform a particular imagery (up or down) for a few seconds. Each trial’s EEG data was labeled accordingly (1 for “Upward” imagery, 0 for “Downward” imagery). Sufficient rest was given between trials to avoid fatigue. Across the session, an equal number of upward and downward trials were recorded per participant to ensure a balanced dataset (for example, 40 trials each, totaling ~80 trials per participant, depending on individual pacing and data quality) [9]. Participants were monitored to minimize eye blinks or muscle artifacts during imagery [20]. The entire experiment typically lasted around 30–45 minutes per participant, including setup, training, and execution [3]. Power Spectral Density (PSD) analysis [16], between the two motor imagery movements are shown as seen in Fig.3.

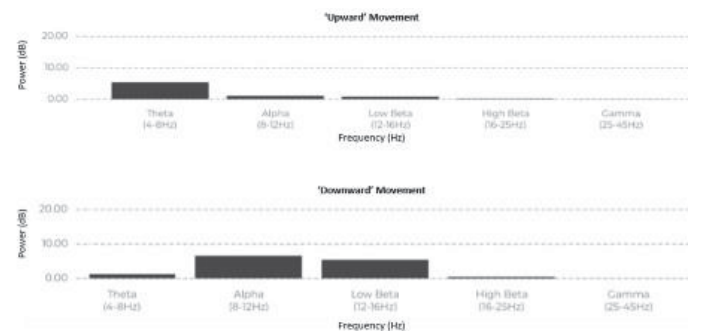


Fig. 3. PSD analysis (frequency-domain differences). Frequency domain power distribution during ‘upward’ and ‘downward’ hand movement

C. Data Acquisition

EEG signals were acquired using the Emotiv Epoc X headset [17], [18], a wireless 14-channel EEG device designed for BCI applications. The electrodes are arranged according to the International 10–20 system as shown in Fig.4 [7], covering frontal, temporal, parietal, and occipital regions. Saline-based wet sensors were placed at locations including AF3/AF4, F3/F4, F7/F8, T7/T8, P7/P8, O1/O2, with two reference electrodes (CMS/DRL) providing a baseline. The device transmits EEG data at 128 Hz sampling rate (with 16-bit resolution), which is sufficient to capture the frequency bands of interest for motor imagery (up to beta range) [12], [13]. During the experiment, the headset was connected to a computer running EmotivPro, which recorded the raw EEG signals from all channels simultaneously as the participant performed the tasks. The recording software also time-stamped and labeled each trial based on the cues, allowing synchronization of EEG segments with the intended command (up or down).

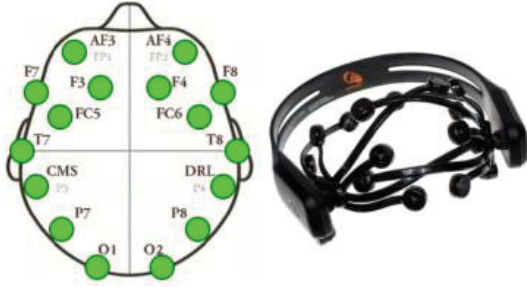


Fig. 4. Emotiv Epoc X electrodes configuration

All recordings were conducted in a quiet lab environment with the participant seated comfortably to reduce movement artifacts. We took basic measures to ensure signal quality: electrode impedances were checked prior to recording, and a stable wireless connection was maintained. The Emotiv Epoc X's built-in filtering and ADC hardware provided baseline noise filtering. Nonetheless, EEG data inherently contained some artifacts from eye blinks, facial muscle activity, and any residual movements. We decided to preserve as much of the raw signal characteristics as possible for analysis rather than aggressively cleaning the data, in order to capture the true challenges of using a consumer-grade EEG in practical settings [8].

D. Data Preprocessing

The raw EEG data from each session were initially processed using Emotiv's software and then exported for offline analysis. We applied basic preprocessing steps uniformly to all participant data. First, a bandpass filter (approx. 0.5–40 Hz) was applied to remove DC drift and high-frequency noise, limiting the analysis to the standard EEG frequency range that includes delta (0.5–4 Hz) through beta (~30 Hz) and low gamma (~40 Hz). Line noise (50/60 Hz) was mitigated via notch filtering if necessary (Emotiv's system includes built-in notch filters) [16]. We kept preprocessing minimal and consistent: no extensive artifact removal (such as independent component analysis) was performed at this stage [9]. The rationale was to retain natural

variability and not overfit the data to any specific participant by manually excluding segments. However, obviously corrupted segments (e.g., if a sensor detached or saturated during a trial) were noted and could be excluded from feature extraction.

Next, we segmented the continuous EEG data into epochs corresponding to each trial. Typically, an epoch covered the duration of the motor imagery command (e.g., a 5–8 second window during which the participant imagined the movement, as prompted) [10]. Each epoch was labelled 0 or 1 according to the commanded class. No baseline correction was applied beyond what the bandpass filtering accomplished, since each trial's mean could be near zero after filtering. We also did not down sample further; the data at 128 Hz was directly used for feature computation [12]. To summarize, preprocessing yielded a collection of labeled EEG epochs per participant, each containing multi-channel time-series data primarily in the 0.5–40 Hz range, ready for feature extraction.

E. Feature Extraction

From each trial's preprocessed EEG segment, we extracted a comprehensive feature vector characterizing the brainwave patterns for that trial. We focused on frequency-domain power features and simple statistical measures that are commonly employed in BCI signal analysis. In particular, band power features were computed for four canonical EEG frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), and beta (12–30 Hz) [10], [12], [5]. For each of the 14 EEG channels, the signal was digitally band-pass filtered to isolate these bands, and the signal power in that band was calculated (averaged over the epoch duration). We define the band-power feature for channel i in band b (for example, F3 in alpha band) as the mean squared amplitude of the EEG signal in that frequency range over time:

$$BP_{i,b} = \frac{1}{N} \sum_{n=1}^N [x_{i,b}(n)]^2,$$

Where $x_{i,b}(n)$ is the EEG sample of channel i after filtering to band b , and N is the total number of samples [16]. This band power is essentially the signal variance in that frequency band (since the EEG is zero-mean after filtering) and reflects the intensity of that band's activity during the task. By extracting delta (δ), theta (θ), alpha (α), and beta (β) power from each of the 14 channels, we obtained an initial set of $14 \times 4 = 56$ features. (The Emotiv Epoc X provides 14 effective EEG channels; the two reference channels are not used as independent features). In addition to band powers, we engineered a number of time-domain features to enrich the feature set. These included the variance of the raw EEG signal on each channel (which, as noted, is closely related to band power but without band-specific filtering), as well as signal energy measures (total power across all frequencies up to 40 Hz for each channel). We also computed composite features such as ratios of power between bands on the same channel (for example, alpha/beta ratio on a channel, which can be informative of engagement vs. relaxation) and across symmetric channels (e.g., difference in beta power between left and right motor cortex areas). These additional features were chosen to capture potential asymmetries and cross-band

interactions relevant to motor imagery. After including these, the total feature vector size was expanded to 84 features per trial. In our final dataset, we had 84 features per trial (each trial corresponding to one instance of either upward or downward imagery) [13].

All features were normalized (z-scored) across the dataset to have zero mean and unit variance prior to classification. This standardization was done using statistics computed from the training data folds only, to avoid any information leakage into the test fold during cross-validation. Feature selection was not aggressively performed at this stage; however, we did keep track of features with extremely low variance or consistently high correlation, as these could be pruned in future iterations to simplify the model (see Results for an analysis of feature variance and correlation). The full feature matrix thus had dimensions of T samples by 84 features, where T is the total number of trials from all participants combined.

F. Classification and Evaluation

We employed a decision tree classifier as the ML model to distinguish between the two mental command classes. A decision tree (using the CART algorithm with the Gini impurity criterion) [19] was chosen for its interpretability and low computational requirements, aligning with the concept of a “weak AI” or lightweight model [21]. This model can highlight which features (channels/bands) are most informative by its branching structure, thus offering insight into neurophysiological relevance in addition to making predictions.

We trained and tested the classifier using a leave-one-participant-out (LOPO) cross-validation scheme to rigorously evaluate how the system performs on an unseen individual [21]. In LOPO, we train the model on data from 9 out of the 10 participants and use the held-out participant’s data exclusively for testing. This process is repeated 10 times, each time with a different participant as the test set, so that we obtain performance metrics for every participant as an independent test case. LOPO is a stringent evaluation because it simulates the real-world scenario of applying a trained BCI model to a new user without any subject-specific recalibration. During each LOPO fold, the decision tree was fitted to the training set (using only the participants in that fold’s training pool). No post-pruning of the tree was performed, as we did not observe severe overfitting given the limited model complexity and small feature set relative to training size. We recorded the predictions of the model on the held-out participant’s trials and compared them to the true labels to compute classification metrics. The primary performance metric was accuracy (the proportion of trials correctly classified). However, given the potential imbalance or differential difficulty in detecting one command vs the other, we also calculated precision and recall for each class, taking the “Upward” command as the positive class for convention. We define these metrics in terms of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) for the positive class (command “Up” = 1) as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN},$$

$$Precision = \frac{TP}{TP + FP},$$

$$Recall = \frac{TP}{TP + FN},$$

These metrics were computed per participant (i.e., per fold) to assess how performance varies among individuals. In each fold’s results, we also identified which class had the higher precision and which class had the lower recall, as an indicator of any bias in the classifier’s predictions for that participant. Because our task is balanced (each participant contributed roughly equal trials of each class), accuracy is a reasonable summary metric; but precision and recall help diagnose whether the classifier tends to favor one command over the other (e.g., always predicting “Up” would yield high recall for “Up” but low precision for it, and poor recall for “Down”). Finally, we aggregated the outcomes from all folds to compute the mean and range of accuracy across participants, and we analyzed instances of particularly high or low performance to glean insights (for example, whether the participants with prior BCI experience fared better, or whether certain EEG patterns correlate with success).

All data processing and classification steps were implemented in Python using libraries such as MNE for EEG filtering and scikit-learn for ML. The use of a simple classifier without extensive parameter tuning was intentional to provide a baseline; more complex models (e.g., deep neural networks) were outside the scope of this initial investigation.

III. RESULTS

A. EEG Feature Extraction

1. Mean and Variance: An overview of the extracted features’ distribution is shown in Fig. 5, which plots the absolute mean and variance of each EEG feature across all trials and participants.

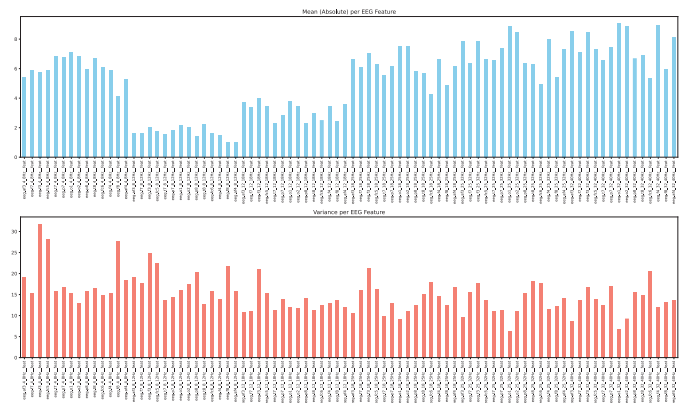


Fig. 5. Variance of EEG band-power features across all participants (P01–P10). (Top) Mean variance values per frequency band across electrodes, highlighting stable and redundant features. (Bottom) Boxplots showing feature-wise variance distributions, with high-variance outliers corresponding to noisy or highly discriminative channels.

Each feature corresponds to a specific electrode–frequency band pair (for band-power features) or a specific statistical measure on a given channel. As seen in the top graph of Fig. 5, features derived from the frontal region in the beta band

(approximately 12–30 Hz, e.g., channels F3-Beta and F4-Beta) exhibited the highest mean values. This suggests that during the motor imagery tasks, these frontal areas produce consistently stronger EEG signals in the beta frequency range, aligning with the known role of the frontal cortex (and motor cortex, which is near F3/F4) in motor planning and execution. In contrast, features from occipital channels in lower frequencies (e.g., O1-Delta, O2-Theta) showed near-zero mean activation, indicating minimal involvement of the occipital (visual) areas in this task, as expected.

The variance analysis (bottom graph of Fig. 5) highlights the stability of each feature. Most features had modest variance, but a few exhibited very low variance (<0.01 across all trials). For instance, certain occipital channel features in delta/theta bands were nearly constant, reinforcing that those carry little information for our motor imagery task. Such low-variance features are likely redundant or noise and could be removed to streamline the classifier with negligible loss of information. On the other hand, extremely high variance was noted in some features (e.g., an electrode–band combination prone to artifacts for one or two participants). While high variance can indicate meaningful dynamic changes, it can also point to inconsistency or contamination (for example, a feature spiking due to muscle artifacts in some trials).

These observations suggest that a future step could involve feature selection: dropping consistently flat features and closely examining features with excessive variance for potential artifact influence. Overall, the mean/variance profiles confirm that the feature set captures known neurophysiological patterns, frontal beta activity for motor tasks, and also includes some channels/frequencies that might be pruned for efficiency.

2. Autocorrelation: To further investigate the temporal dynamics of the EEG features, Fig. 6 presents the autocorrelation function for selected features (each subplot corresponding to one feature), computed across time lags within each trial.

Autocorrelation measures how similar a signal is to itself after a time shift, and it reveals the presence of rhythmic or persistent patterns in the data. As expected, all features exhibit a strong autocorrelation peak at lag 0 (correlation = 1 by definition). Beyond lag 0, we observed two general patterns. For many features, especially those from sensorimotor-related channels (e.g., electrodes over parietal cortex) in the alpha and beta bands, the autocorrelation decays gradually over a range of 50–200 ms before dropping to near zero. This gradual decay suggests the presence of oscillatory neural rhythms (like the mu rhythm around 10 Hz or beta oscillations ~20 Hz) which impart temporal structure to the signal. Such structure is expected during MI, as the brain engages oscillatory sensorimotor patterns. In contrast, a subset of features (often those from channels or bands not strongly engaged by the task) showed a very rapid drop-off in autocorrelation, essentially becoming uncorrelated at even small lags. These flat autocorrelation profiles imply that the feature is dominated by broadband noise or random fluctuations rather than a repeatable pattern. For example, an electrode over occipital cortex during an

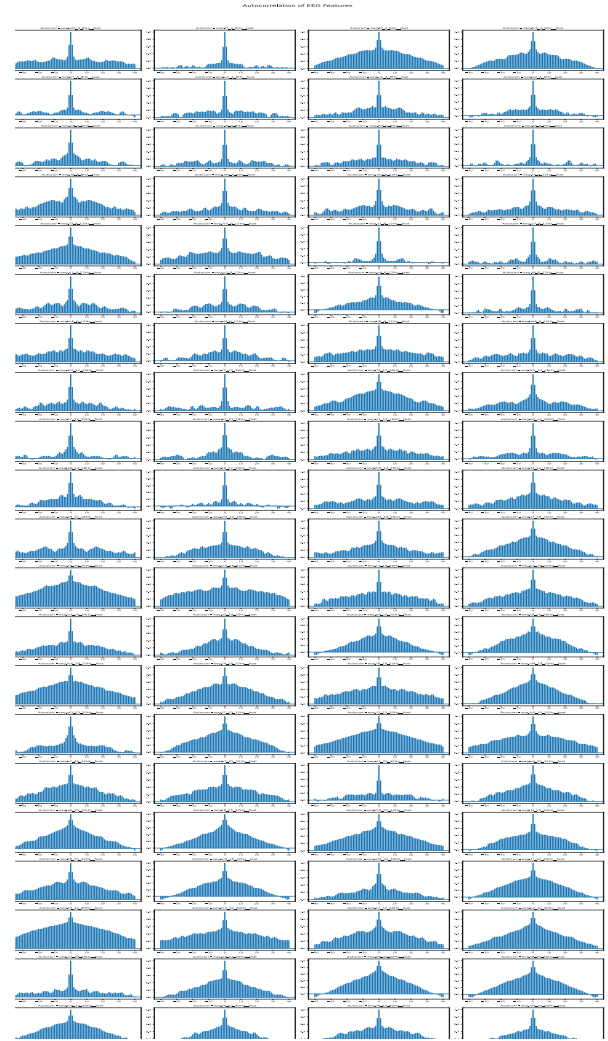


Fig. 6. Autocorrelation plots of EEG features. (Top) Beta-band activity over motor cortex electrodes (C3, C4) shows sustained rhythmicity during motor imagery. (Bottom) Theta-band activity demonstrates faster decay, reflecting less stable oscillatory structure across participants.

imagination of hand movement might show quickly decorrelating activity, since it's mostly noise relative to the task.

Importantly, none of the features displayed pathological autocorrelation (e.g., sustained high correlation at large lags that could indicate unremoved trends). The features that showed structured autocorrelation (gradual decay) likely correspond to meaningful neural signals (such as ongoing alpha/ beta rhythms modulated by MI), whereas those that were almost delta-correlated (spiky at lag 0 and nearly zero afterwards) might be candidates for exclusion or down-weighting in classification. Preserving features with richer autocorrelation structure is beneficial because they reflect the dynamic nature of the brain's response to motor imagery. In summary, the autocorrelation analysis confirms that a significant portion of our features capture temporally structured EEG activity consistent with sensorimotor rhythms, whereas others behave like noise. This justifies efforts to optimize the feature set by focusing on features that carry temporal information relevant to the task.

3. Feature Correlations: We examined pairwise correlations between all features to identify redundant features and interdependencies. Fig. 7 depicts the Pearson correlation matrix (heatmap) for the full feature set.

Several interesting patterns emerge. We observe distinct clusters of high correlation among features that are physiologically related. For example, features from symmetric electrode sites on opposite hemispheres (such as F3 and F4, or T7 and T8) often show strong positive correlation across trials. This implies that when, say, left-frontal beta power increases during a trial, right-frontal beta power tends to also increase, which is reasonable given bilateral cortical involvement in imagined movement. Similarly, features from the same electrode but adjacent frequency bands are correlated – e.g., at electrode P8, the alpha-band power and beta-band power have moderate correlation, possibly reflecting a general level of arousal or engagement affecting both bands simultaneously. These correlations suggest some redundancy: a model might not need both features if one can predict the other. We also note that certain features are virtually uncorrelated (correlation near 0) with most others, indicating they might carry unique information (for instance, a feature in delta band might capture a different aspect of the signal than features in higher bands).

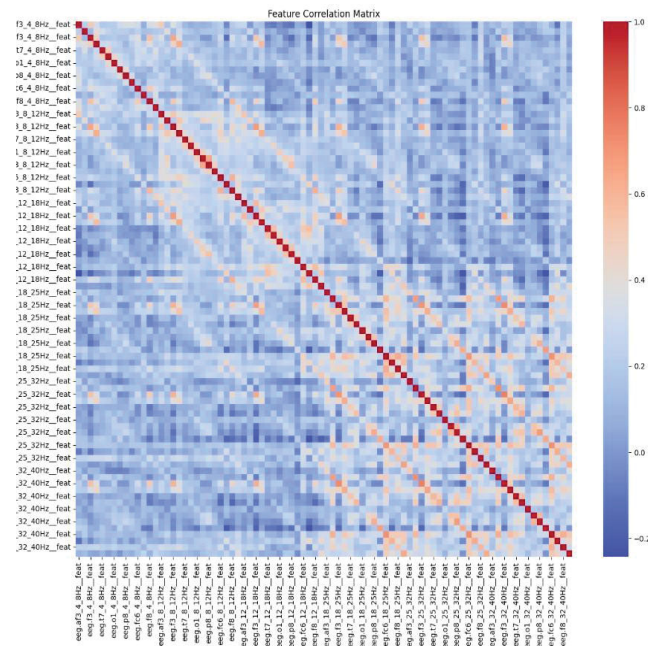


Fig. 7. Correlation matrix of EEG band-power features. (Left) Correlations across frontal and motor cortex electrodes. (Right) High correlations (>0.9) between symmetric electrode pairs (e.g., F3–F4, C3–C4) indicating redundancy and potential for dimensionality reduction.

In a few cases, we found very high correlation ($r > 0.9$) between specific feature pairs. For instance, the pair F3-Theta and F4-Theta had an almost identical response pattern across all trials, as did O1-Alpha and O2-Alpha. Such highly redundant features introduce multicollinearity, which can unnecessarily complicate the classifier and possibly reduce generalization if the model tries to give weight to both. In future work, one could remove or merge one of each highly correlated pair (using techniques like principal component analysis or just dropping

one). On the positive side, the presence of low or zero-correlation pairs (e.g., a frontal beta feature vs. an occipital theta feature) suggests our feature set includes complementary information streams. This is advantageous for classification, as combining uncorrelated features can improve the classifier's ability to distinguish classes. The correlation matrix thus provides a map of feature space, indicating which features cluster together. In summary, Fig. 3 indicates that while many features are interrelated (especially symmetric channels and close frequency bands), there are also sufficiently independent features that capture different facets of the EEG data. Managing the few strongly redundant features ($r > 0.9$) could streamline the model, but otherwise the diversity of features is likely beneficial.

4. Feature Distributions: We also visualized the distribution of feature values to identify outliers and gauge overall signal quality. Fig. 8 shows horizontal boxplots for a selection of features (for brevity, not all 84 are shown in one figure, but representative ones across different channels and bands). Each boxplot summarizes the feature's values over all trials (pooled from all participants). The median, interquartile range, and any extreme points are indicated. We found that for most features, the central 50% of values fell in a plausible physiological range (e.g., EEG band power values that are consistent with typical EEG magnitudes). However, some features exhibited significant outliers. For example, Betaband power at electrode F8 showed several trials with values far above the upper quartile, and Alpha band power at O2 had a few unusually low points. Investigating these revealed that such outliers often corresponded to specific participants or trials where transient artifacts occurred – e.g., a muscle tension or jaw movement causing a burst in high-frequency power at F8, or momentary signal dropouts at O2 leading to abnormally low power. In general, frontal and temporal sites tended to have more high-value outliers (likely muscle-related), while occipital sites had occasional low outliers (possibly due to momentary disconnections). Crucially, the presence of outliers did not skew the classifier training unduly, because the decision tree is relatively robust to monotonic transformations and outliers (it splits ranges based on ordering). Nevertheless, these outliers flag opportunities for improved preprocessing: techniques like artifact subspace removal or trial rejection could be applied to handle extreme cases. Notably, none of the features had a median that was zero or extreme, confirming that our earlier filtering steps (bandpass, etc.) kept the data within sensible ranges. The spread of most boxplots is moderate, indicating that while there is subject-to-subject variability, the feature ranges overlap significantly between participants – a condition that at least allows a single model to attempt generalization. In summary, Fig. 8 suggests that the overall feature distributions are reasonable, with only a minority of trials showing aberrant values. Outliers, where present, are explainable by artifacts, and their effect might be mitigated by robust statistics or trimming in future work. The median differences between features also reiterate which features have higher typical values (e.g., frontal beta features) versus lower (occipital delta), complementing the mean analysis.

B. Classification Performance (LOPO Cross-Validation)

We evaluated the classification of “Upward vs. Downward” motor imagery for each participant using the LOPO scheme described earlier. The overall accuracy of the system varied substantially across individuals. It ranged from a low of 39% (for the worst-case participant, which is below the 50% chance level for a binary task) up to a high of 66% (for the best-performing participant) [9], [13], [19]. The mean accuracy across the 10 participants was 55.5%, which is only slightly above chance, reflecting the challenge of building a one-size-fits-all model for EEG patterns.

Detailed performance metrics for each participant are summarized in Table II. We report each participant’s accuracy, as well as the precision and recall for the better-classified class (“Best Class Precision”) and the weaker class (“Weak Class Recall”), to illustrate any bias in the classifier’s decisions.

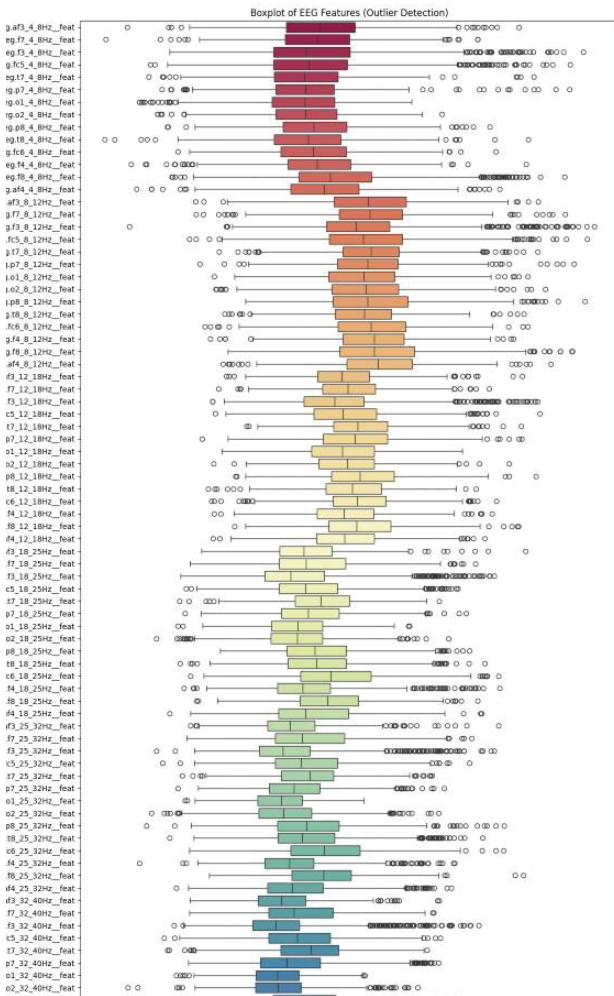
TABLE II. LOPO AND DECISION TREE SUMMARY

No.	Accuracy	Best Class Precision	Weak Class Recall	Notes
P01	0.48	0.48 (Class-1)	0.28 (Class-1)	Low Separability
P02	0.39	0.41 (Class-1)	0.26 (Class-1)	Difficult Sample; Likely Noisy
P03	0.55	0.60 (Class-0)	0.24 (Class-1)	High False Negatives
P04	0.52	0.51 (Class-1)	0.35 (Class-0)	Inverse Balance in Performance
P05	0.49	0.50 (Class-1)	0.33 (Class-0)	Struggles With 0→1 Confusion
P06	0.62	0.65 (Class-1)	0.49 (Class-1)	Good Class Balance
P07	0.64	0.69 (Class-0)	0.56 (Class-1)	Best Generalization So Far
P08	0.55	0.58 (Class-0)	0.53 (Class-1)	Balanced But Average Performance
P09	0.66	0.69 (Class-1)	0.63 (Class-1)	Best Classifier Performance
P10	0.65	0.69 (Class-1)	0.56 (Class-1)	Excellent Participant-Wise Balance

Notably, two participants (P09 and P10) achieved the highest accuracies, 66% and 65% respectively. These participants’ models not only got more trials correct overall, but also exhibited relatively balanced precision and recall between the two command classes (for instance, P09’s precision and recall for the target class were 0.69 and 0.63, both fairly high – indicating the model was good at detecting both “Up” and “Down” for them). This suggests that their EEG signals during the two imagery tasks were more separable and consistent.

In contrast, the participant with the lowest accuracy (P02, 39%) showed a strong bias: the model often misclassified one of the classes for them. Specifically, P02’s precision for the “Downward” class was only 0.41 and the recall for “Upward” was 0.26, meaning the classifier frequently confused their Downward imagery as Upward (many false positives for Up) and also missed many of their Up trials (false negatives). This points to low separability in their EEG patterns – P02’s data might be an example of a participant for whom the EEG signals did not show clear distinctions between the two mental commands (possibly an instance of BCI illiteracy or simply poor signal quality that day).

Other participants fell in between these extremes. For example, P07 reached 64% accuracy, with a slight edge in detecting class 0 (“Down”) well (precision 0.69 for class 0) but also decent recall for that class (0.56). P04 and P05 were



around 50% accuracy, indicating the classifier was barely above random guessing for them. In the case of P04, the precision for “Upward” was ~ 0.51 but recall for “Down” was only 0.35, suggesting the model tended to over-predict the Upward class (bias toward class 1).

Looking across all participants, there is a clear trend of high inter-subject variability. Some participants’ brain signal patterns generalized well to the group-derived model (those with $>60\%$ accuracy, such as P06, P07, P09, and P10), whereas others did not. This variability underscores the known phenomenon that a significant portion of users have difficulty achieving reliable control in EEG-based BCIs. In our case, all participants were able to complete the tasks, but clearly the ML model could not accommodate the idiosyncrasies of each person’s EEG within this simple feature space.

To gain insight, we note that participants such as P01 and P02 (lowest accuracies) were characterized by low separability and noisy data, meaning the classifier likely found no clear rule to distinguish their two states. Participants in the middle range (e.g., P03, P04, P05, P08) showed specific issues such as high false negatives or imbalance in performance (favoring one class), which could be due to those individuals consistently performing one imagery task more strongly than the other or the model being biased. The best performers (P06, P07, P09, P10) are noted as having good class balance and generalization, which implies their patterns were not only distinct, but also somewhat representative of the group – the classifier trained on others was still effective on them, hinting that their brain responses to MI were “typical” relative to the population data.

In terms of precision and recall aggregated across all participants, the model’s precision for the positive class (“Upward” imagery) averaged about 0.55, and recall about 0.46, indicating a slight tendency to over-predict the downward (class 0) for some participants (missing upward trials). However, these numbers varied widely person to person. No significant systematic bias toward one command was observed when averaging across all – the inconsistencies largely canceled out.

In summary, the classification results demonstrate proof-of-concept success in some participants and difficulties in others. When a generic model is applied, roughly half the participants achieved around 60+% accuracy, whereas the rest hovered near chance. This clearly indicates that a one-model-for-all approach is suboptimal in EEG BCIs and that personalized calibration is likely necessary to reach high control accuracy for each user. Nonetheless, the fact that a simple decision tree on basic features could exceed chance for a majority of participants (and reach mid-60s for the best) is encouraging; it suggests that there is a discernible EEG difference between imagining upward vs. downward hand movement, even if subtle and user-dependent. The performance might be sufficient for a rudimentary two-command BCI if combined with error mitigation strategies (for example, using a “confirm” step or majority voting over multiple predictions to execute a drone command only when the classifier is confident).

IV. DISCUSSION

A. Comparison with Prior Work

The concept of controlling UAVs via BCIs is not new; prior research has explored different modalities. For example, steady-state visual evoked potential (SSVEP)-based systems have achieved multiclass drone navigation with accuracies often exceeding 90% in controlled laboratory conditions [6], [24], while P300-based systems have also been used reliably for selection tasks [4]. These approaches, however, typically require external stimuli and individual calibration.

Our work instead employed MI, which is more challenging due to the subtlety of EEG changes and the phenomenon of BCI “inefficiency,” wherein some users cannot produce distinct MI signals [10], [9]. Prior MI-based UAV studies have reported varying levels of success. For example, Vijayendra et al. demonstrated high offline classification precision using a 14-channel consumer EEG device [12], combined with more advanced pre-processing and neural network models.

In comparison, the present study used only a lightweight decision tree model [19], [21]. The highest accuracy observed was 66% for P09, closely followed by P10 at 65%. These accuracies were achieved without participant-specific calibration, underscoring both the feasibility and limitations of using generic models. Typically, within-subject MI classification can reach 70–80% after training [20], [23], but across-subject generalization often falls closer to chance level unless advanced techniques such as transfer learning are applied.

B. Inter-Subject Variability

The A key observation in this study is the large inter-subject variability [9][8]. Performance ranged from 39% (P02) to 66% (P09). Such variability aligns with findings in BCI literature, where it is estimated that 15–30% of users struggle to achieve reliable control, a phenomenon referred to as “BCI illiteracy” [9], [7].

Participants P06, P07, P09, and P10 achieved relatively high and balanced accuracies (above 60%), suggesting that their EEG signals showed consistent separability between upward and downward imagery. In contrast, P02 achieved the lowest accuracy, with the classifier frequently misclassifying their “Downward” imagery as “Upward.” This result may reflect poor signal quality, inconsistent performance of the task, or inherently lower separability of neural patterns.

Participants in the middle range (P01, P03, P04, P05, P08) often showed systematic biases, such as high false negatives in one class or a tendency toward over-predicting the “Upward” class. For example, P04 exhibited bias toward class 1 (“Upward”), while P03 showed high false negatives for the same class. These issues highlight the idiosyncratic nature of EEG signals and reinforce the need for individualized model adaptation.

C. Rationale for Binary MI-based Control

This investigation focused deliberately on a binary classification problem: distinguishing imagined upward versus downward movements [3]. This choice was motivated both by practical and theoretical considerations. From a practical perspective, starting with two commands simplifies training and evaluation, and many UAV control scenarios can be reduced to binary decisions (e.g., hover vs. land). From a neuroscience perspective, MI of distinct movements can engage different cortical areas and rhythms [10], [11]. Although right-hand upward vs. downward imagery may produce less distinct neural signatures than left- vs. right-hand imagery, the semantic mapping to UAV control tasks (up and down) was intuitive. Prior demonstrations, such as [3], have shown that binary MI can be sufficient for basic UAV commands, even if accuracy is modest. Our results, with several participants achieving above 60%, support this rationale.

D. Implications for BCI-Controlled IoT Devices

The results show that consumer-grade EEG hardware (Emotiv Epoc X) can capture signals enabling above-chance binary classification of motor imagery. This is promising for integration of BCIs into IoT systems, including UAVs, smart wheelchairs, or simple smart home commands [26]. However, accuracy levels around 60–65% are insufficient for standalone control in safety-critical contexts. Instead, BCIs may serve as a supervisory layer, providing high-level inputs while autonomous systems manage low-level navigation or stabilization. For example, a UAV could execute “land” when the BCI classifier is highly confident, while uncertain outputs could be validated by secondary modalities (voice, manual override). Another implication concerns user training. Some participants, such as P02, may improve significantly with repeated practice, feedback, or adaptive algorithms. Studies show that BCI performance can increase over time as users learn to produce more consistent neural patterns. Co-adaptive approaches, in which the system adjusts to the user while the user learns, may further improve control [27].

E. Limitations and Future Work

Several limitations must be acknowledged. First, the sample size was small (10 participants), with limited demographic diversity. Second, data were collected in a single session without participant-specific calibration, reducing performance potential. Third, the system was evaluated offline rather than in real-time on an actual UAV. Finally, only two commands were considered, limiting functional scope.

Future research directions include:

1. Personalized Calibration: Even short calibration sessions could adapt the model to each user’s EEG [7], [8], [9].
2. Advanced Classification Methods: Deep learning and Riemannian geometry-based approaches have shown superior MI performance in competitions [21], [23].

3. Hybrid BCI Paradigms: Combining MI with P300 or SSVEP could increase reliability and enable confirmation mechanisms [4], [6], [14], [24].
4. Real-Time Implementation: Testing with an actual UAV to assess information transfer rate, latency, and user experience [3], [12], [13], [25].

Despite these limitations, the fact that participants like P09 and P10 achieved above 65% accuracy with no calibration demonstrates proof-of-concept feasibility. With personalization, performance could be improved, paving the way for practical BCI–IoT integration.

IV. CONCLUSION

In this paper, we presented a systematic exploration of an EEG-based BCI for controlling UAV using binary motor imagery tasks. Our approach used a consumer-grade 14-channel Emotiv Epoc X headset, lightweight feature extraction, and a decision tree classifier.

The results demonstrated proof-of-concept feasibility. Several participants, such as P09 and P10, achieved accuracies above 65% without calibration, showing that separable EEG patterns for imagined upward and downward movements can be detected even with a simple classifier. However, performance varied significantly across individuals: P02 achieved only 39%, while most participants clustered around 50–60%. This inter-subject variability highlights the limitations of applying a one-model-fits-all approach to EEG-based BCI systems.

Despite modest average accuracy (55.5%), the study demonstrates that neural interfaces can provide meaningful binary control signals for UAV supervision. A binary BCI may not yet replace manual controllers, but it can serve as a useful complementary modality — for example, providing a “hover” or “land” command in safety-critical contexts or serving as an assistive option for users with motor impairments.

The findings also emphasize the importance of subject-specific calibration, adaptive algorithms, and potentially hybrid paradigms to improve performance. Future work should focus on personalization, real-time implementation, and extension beyond two commands. With these improvements, BCI-based UAV control could become a practical tool in Internet of Things (IoT) environments, bridging human thought and machine action.

ACKNOWLEDGMENT

The authors would like to thank all participants for their involvement in the experiments and their commitment during the training and recording sessions. Special thanks to Saint Petersburg State Electrotechnical University “LETI” for providing access to laboratory facilities and equipment. This research of authors M.A.G., I.I.V., A.A., and G.R.I., was funded by the Ministry of Science and Higher Education of the Russian Science Foundation (Project “Goszadanie”, FSEE-2024-0011).

REFERENCES

- [1] M. A. Lebedev and M. A. Nicolelis, "Brain-machine interfaces: past, present and future," *Trends in Neurosciences*, vol. 29, no. 9, pp. 536–546, 2006. doi:10.1016/j.tins.2006.07.004
- [2] J. R. Wolpaw and E. W. Wolpaw, *Brain-computer interfaces: Principles and practice*. Oxford University Press, 2020.
- [3] LaFleur, K., Cassady, K., Doud, A., Shades, K., Rogin, E., & He, B. (2013). Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain-computer interface. *Journal of Neural Engineering*, 10(4), 046003
- [4] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials (P300)," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523.
- [5] J. Lee, S. Kim, and H. Park, "Analyzing beta and theta EEG patterns for enhanced motor imagery recognition in brain-computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, no. 5, pp. 678–686, 2022.
- [6] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *Journal of Neural Engineering*, vol. 6, no. 4, 046002, 2009.
- [7] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces: A review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2020.
- [8] R. Abiri, S. Borhani, E. W. Sellers, Y. Jiang, and X. Zhao, "A comprehensive review of EEG-based brain-computer interface paradigms," *Journal of Neural Engineering*, vol. 16, no. 1, 011001, 2019.
- [9] C. Sannelli, C. Vidaurre, K.-R. Müller, and B. Blankertz, "A large-scale screening study with a SMR-based BCI: Categorization of BCI users and differences in their SMR activity," *PLOS ONE*, vol. 14, no. 1, e0207351, 2019.
- [10] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE*, vol. 89, no. 7, pp. 1123–1134, 2001.
- [11] G. Pfurtscheller, G. R. Müller-Putz, R. Scherer, and C. Neuper, "Rehabilitation with brain-computer interface systems," *Computer*, vol. 41, no. 10, pp. 58–65, 2008.
- [12] A. Vijayendra, S. K. Saksena, R. M. Vishwanath, and S. N. Omkar, "A performance study of 14-channel and 5-channel EEG systems for real-time control of UAVs," in *Proc. IEEE Int. Conf. Robotic Computing (IRC)*, 2018, pp. 183–188.
- [13] B. H. Kim, M. Kim, and S. Jo, "Quadcopter flight control using a low-cost hybrid interface with EEG-based classification and eye tracking," *Computers in Biology and Medicine*, vol. 51, pp. 82–92, 2014.
- [14] S.-J. Kim, S. Lee, H. Kang, S. Kim, and M. Ahn, "P300 brain-computer interface-based drone control in virtual and augmented reality," *Sensors*, vol. 21, no. 17, 5765, 2021.
- [15] J.-A. Cervantes et al., "COGNIDRON-EEG: A system based on a brain-computer interface and a drone for cognitive training," *Cognitive Systems Research*, vol. 78, pp. 48–56, 2023
- [16] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.
- [17] EMOTIV Inc., "EPOC X – 14 Channel Wireless EEG Headset (product page).
- [18] EMOTIV, "EPOC X User Manual – Technical Specifications" 2020.
- [19] A. Gramfort et al., "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, 267, 2013.
- [20] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Wadsworth, 1984.
- [21] F. Lotte et al., "A review of classification algorithms for EEG-based brain-computer interfaces: a 10-year update," *Journal of Neural Engineering*, vol. 15, no. 3, 031005, 2018.
- [22] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [23] V. J. Lawhern et al., "EEGNet: A compact convolutional network for EEG-based brain-computer interfaces," *arXiv:1611.08024*, 2016 (and later journal versions).
- [24] A. Chiuzbaian, J. Jakobsen, and S. Puthusserypady, "Mind controlled drone: An innovative multiclass SSVEP-based BCI," in *Proc. 7th Int. Winter Conf. on Brain-Computer Interface (BCI)*, 2019, pp. 1–5.
- [25] D.-H. Lee, J.-H. Jeong, H.-J. Ahn, and S.-W. Lee, "Design of an EEG-based drone swarm control system using endogenous BCI paradigms," *arXiv:2012.03507*, 2020
- [26] J. Peksa et al., "State-of-the-Art on Brain-Computer Interface Technology," *Sensors*, 23(14):6428, 2023.
- [27] F. Atilla et al., "Gamification of motor imagery BCI training: A review," *Brain-Computer Interfaces*, 2024.