Human Trust-Driven Adaptive Control for Unmanned Aerial Swarms

Jesus A. Orozco, Coleman Walsh, and Panagiotis Artemiadis*, *IEEE Senior Member*
Mechanical Engineering Department
University of Delaware, Newark, DE 19716, USA.
jaon@udel.edu, cowalsh@udel.edu, partem@udel.edu
**Corresponding Author*

***Abstract*—This paper introduces an innovative Human-Swarm Interaction (HSI) architecture leveraging a Brain-Computer Interface (BCI) to evaluate operator trust in real-time during collaborative multi-UAV navigation. While Electroencephalography (EEG)-based trust detection is well-documented, its implementation remains predominantly restricted to offline assessments. Addressing the critical need for online, adaptive control in aerial robotics, we propose a k-Nearest Neighbors (k-NN) classifier that differentiates trusting from distrusting cognitive states with over 90% accuracy. Integrated into a closed-loop BCI, this model dynamically adjusts aerial swarm formations. Experimental validations demonstrate the system's capacity to swiftly identify operator distrust, prompted by sub-optimal UAV behaviors, and autonomously revert the swarm to a stable, trusted configuration. These findings highlight that higher-order human factors, particularly trust, can effectively govern adaptive controllers in unmanned aerial systems, facilitating hands-free operations for complex missions.**

I. Introduction

Unmanned Aerial Vehicle (UAV) swarms provide critical advantages in search-and-rescue, surveillance, and defense operations due to their inherent scalability, fault tolerance, and adaptability [1], [2]. These bio-inspired, decentralized systems [3] transition human operators from continuous manual control to high-level supervision, significantly reducing cognitive load. However, the successful deployment of autonomous multi-UAV teams depends heavily on effective Human-Swarm Interaction (HSI). To foster operator trust, aerial swarms must be perceived as highly responsive to human intent. This supervisory dynamic introduces unique challenges, necessitating bi-directional transparency to prevent operational failures [4], [5], [6]. Consequently, developing intuitive, hands-free communication modalities is paramount for enabling seamless, adaptive collaboration between human operators and aerial collectives in complex environments.

Brain-Computer Interfaces (BCIs) offer a promising hands-free modality for aerial swarm control. Typically utilizing non-invasive Electroencephalography (EEG) and subject-specific machine learning models, BCIs decode neural activity to execute targeted tasks. Historically, BCI applications in Human-Swarm Interaction (HSI) have prioritized explicit, direct control over the collective. For instance, hybrid systems combining physical inputs with motor imagery [7], or utilizing speech and motor imagery [8], [9], have been employed to directly manipulate swarm topology and kinematics. However, relying solely on explicit commands remains cognitively demanding for drone operators. Shifting toward implicit cognitive constructs, such as operator trust, provides a more intuitive and naturalistic framework for human-UAV teaming, enabling adaptive multi-agent behaviors

Trust is a fundamental component of effective collaboration across diverse operational settings. However, artificial agents, such as autonomous multi-UAV systems, inherently lack the capacity to directly perceive a human operator's cognitive state and, consequently, cannot autonomously adapt to emergent distrust. In this context, we adopt the widely accepted definition of trust as "the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability" [10]. Within the domain of HSI, trust evaluation is broadly categorized into indirect and direct approaches. Indirect methodologies rely on predictive models driven by external variables, such as system dependability, operator demographics, interface transparency, and swarm status metrics [11], [12], [13], [14], [15]. While promising for adaptive control, these indirect models often struggle with robustness when confronted with dynamic, unmodeled operational shifts, potentially compromising human-UAV teaming. Conversely, direct approaches utilizing EEG to decode operator trust have successfully yielded classification models with approximately 90% predictive accuracy [16], [17]. Despite these advances in directly capturing trust-related neural dynamics, a critical question remains: can operators leverage these implicit cognitive states to modulate an aerial swarm's behavior in real time?

To address this question, this work introduces a closed-loop BCI architecture that directly influences aerial swarm behaviors based on the operator's perceived trust.

Building upon foundational studies that established the feasibility of decoding human trust during swarm monitoring [16], [17], we advance the state-of-the-art by transitioning from passive, offline assessment to active, online control. Utilizing the signal processing pipelines established in [17], we train a classifier capable of discriminating between high- and low-trust states with approximately 93% accuracy. Upon deployment within a real-time framework, the system success-

This material is based upon work supported by the National Science Foundation under Grant No. #2014264, as well as by the U.S. Air Force Office of Scientific Research (AFOSR) award FA9550-18-1-0221.

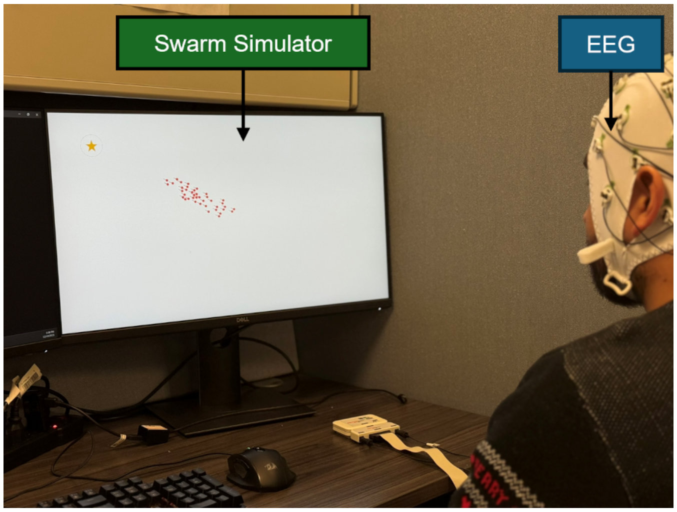

Fig. 1. Experimental setup: A human operator equipped with an EEG cap interacts with a simulated multi-UAV swarm via a closed-loop Brain-Computer Interface.

fully detects operator distrust, often triggered by suboptimal or unpredictable UAV dynamics, and autonomously commands the collective to revert to a secure, desirable formation. To validate our neural decoding, these physiological predictions are rigorously cross-referenced against standardized Human-Robot Interaction (HRI) trust scales [18]. Furthermore, we assess the cognitive load imposed by the BCI to evaluate operator mental demand. Ultimately, these contributions establish a robust, hands-free paradigm for adaptive human-machine teaming, directly applicable to high-stakes aerial missions such as search-and-rescue and defense.

II. Methods

Building upon evidence that human trust is neurologically discernible during both passive and dynamic swarm observations [16], [17], we developed a virtual simulation where an operator interacts with a multi-UAV team via a trust-responsive BCI.

The subsequent sections detail the training data, the aerial swarm simulator, and the EEG-driven BCI architecture utilized for real-time human-swarm interaction.

A. Experimental Setup

During the passive experimental phase, 32-channel EEG data and event markers were recorded at 500 Hz. Signals were acquired using a wireless LiveAmp 32 system equipped with actiCAP slim active electrodes and an actiCAP snap head cap (Brain Products GmbH, Gilching, Germany).

Electrodes were positioned according to the International 10-20 system [19] to effectively capture the operator's neural responses to the multi-UAV simulation (see Fig. 1).

B. Task Description

The simulation environment featured 40 virtual UAVs, visualized in a top-down 2D workspace, tasked with navigating to a diagonally opposite target within a strict 25-second timeframe. The aerial agents initialized from a designated corner with random headings. During the system's training phase, operators passively observed the UAV collective, continuously evaluating its trustworthiness. In the subsequent closed-loop BCI phase, operators could actively toggle the swarm's navigational mode by eliciting a distrustful cognitive state when observing undesirable kinematics.

We implemented two distinct multi-UAV behavioral modes adapted from [16]: *Surround* and *Explore*. The *Explore* mode generated stochastic, diffusive flight paths driven solely by inter-UAV repulsion forces. Conversely, the *Surround* mode integrated Reynolds' standard flocking principles (separation, alignment, and cohesion) [3] with an attractive goal vector to ensure stable, coordinated navigation toward the target. Both algorithms utilized boundary repulsion to confine the UAVs within the operational airspace [17]. Functionally, the goal-directed *Surround* behavior reliably completed the mission within the time limit, whereas the untrustworthy *Explore* behavior consistently failed to meet the temporal constraints.

C. Swarm Simulator

To generate and manage the virtual multi-UAV environment and the agents within it, we utilized the Unity game engine [20]. This represents a significant architectural upgrade from the Python-based tkinter framework previously employed in [16], specifically chosen to facilitate seamless scene transitions and the integration of the real-time BCI framework. For accurate data collection, both the recorded EEG signals and the corresponding experimental event markers were rigorously time-synchronized via the Lab Streaming Layer (LSL) Recorder software [21]. The visual interface of the simulation presents an unobstructed, open-air workspace where solely the aerial agents and their designated target are visible to the operator. To provide clear visual distinction between the two navigational modes, the UAVs executing the goal-oriented *Surround* behavior are rendered in red, whereas those operating under the stochastic *Explore* behavior are depicted in green. Through the integrated BCI system, the simulation empowers the operator to actively manipulate the active swarming behavior, effectively overriding the collective's default operational mode based on real-time neural inputs. Upon successfully navigating to the target coordinate, an individual UAV agent is taken offline and subsequently removed from the visual display. A visual depiction of the simulated multi-UAV environment, exactly as observed by the participant during an experimental trial, is provided in Fig. 2.

D. Experiment Details

A total cohort of five male participants ($N = 5$), possessing an average age of 29.2±2.9 years, was recruited for this study. Each individual satisfied all predefined inclusion criteria before commencing the experimental procedures. The research protocol received formal approval from the University of Delaware Institutional Review Board (IRB ID: 1487701-9), and documented informed consent was obtained from every subject prior to participation.

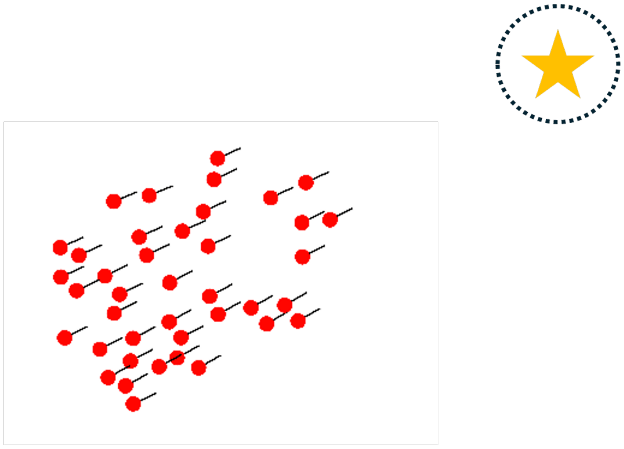

Fig. 2. Visual representation of the simulated multi-UAV environment. The virtual aerial agents (depicted as red circles) navigate across an unobstructed workspace while executing the *Surround* swarming behavior, advancing toward a designated target visually marked by a dashed circle containing a concentric yellow star.

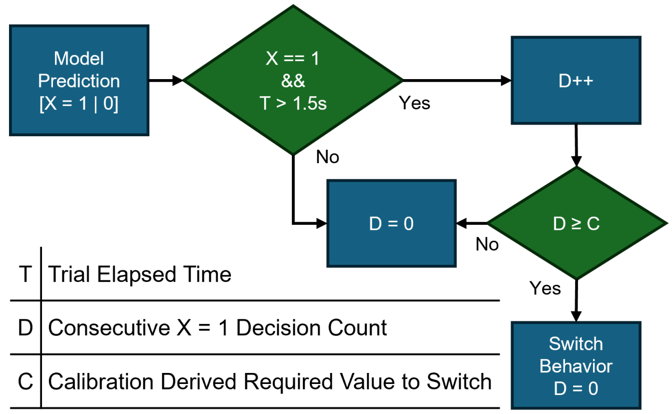

Fig. 3. Logical flowchart of the BCI controller governing multi-UAV behavior transitions. Output values X = 1 and X = 0 denote distrustful and trusting cognitive states, respectively. Following a 1.5-second startup delay to clear pre-trial predictions, the system toggles swarm behavior when consecutive X = 1 states reach calibrated thresholds. Any trusting state (X = 0) resets this counter.

Before initiating the experimental sessions, participants received comprehensive instructions outlining their required tasks, mirroring the protocol established in [16]. Operators were instructed to continuously evaluate the perceived trustworthiness of the aerial collective's ability to successfully execute its navigational objectives throughout the duration of every trial. To mitigate cognitive fatigue, a 5-second blank screen was presented immediately preceding each simulation run, affording the subject a momentary resting period. The experimental design comprised 20 distinct trials in total: 10 dedicated to each specific multi-UAV swarming modality, further divided into 5 iterations for each of the two starting locations.

Following the completion of this foundational trust evaluation stage, operators were granted a scheduled break of roughly 10 minutes. Throughout this designated pause, the newly acquired EEG signals underwent computational processing to calibrate and train the subject-specific k-Nearest Neighbors (k-NN) classifier. Simultaneously, the participants filled out a standardized trust perception questionnaire [18] to quantify their subjective assessment of the distinct UAV behaviors witnessed during the preceding observation period.

The subsequent, active BCI-driven phase exactly mirrored the environmental parameters and strict time limitations of the initial training stage. The underlying control framework, visually depicted in Fig. 3, dictated how the simulated aerial swarm adapted to the human operator's decoded cognitive state. For a successful outcome in the *Surround* scenarios, subjects needed to sustain a neurological state aligning with the baseline of high trust formulated during the training block. On the other hand, successfully navigating the *Explore* scenarios required the subject to initially manifest a state of distrust, prompted by the collective's erratic and undesirable kinematics. This detected distrust would trigger the BCI controller to autonomously rectify the UAVs' flight paths, subsequently allowing the operator's mental state to naturally shift back to one of trust. After concluding these active control sessions, subjects utilized a digital application to complete the NASA Task Load Index (NASA-TLX) assessment [22], providing a quantitative measure of the cognitive demands imposed by operating the trust-based neural interface.

E. Data Analysis For Trust Assessment Phase

The entirety of the recorded neural data underwent processing and analysis within the MATLAB computational environment, utilizing the open-source EEGLAB toolbox [23]. We adopted a windowed data segmentation pipeline that closely mirrors the dynamic methodology outlined in [17]. However, for the purposes of this multi-UAV application, feature extraction was exclusively constrained to power spectral density metrics, aligning perfectly with the passive analytical approach previously validated in [16]. The processing steps are outlined below.

1) EEG Data Processing: The acquired raw EEG signals were first subjected to temporal frequency filtering, employing high-pass and low-pass filters established at cutoff frequencies of 1 Hz and 40 Hz, respectively, to isolate the relevant neurological bands.

Subsequent to this initial filtering, the data streams were further refined within the EEGLAB environment utilizing the Artifact Subspace Reconstruction (ASR) algorithm [24] to systematically identify and mitigate transient noise and non-brain artifacts. Upon completion of the ASR cleaning procedure, the entire neural dataset was re-referenced to the common average computed across all active channels. Ultimately, the continuous recording was segmented into discrete epochs based on the synchronized trial initiation and termination event markers. This precise temporal windowing allowed us to accurately map the isolated neural data to the corresponding multi-UAV operational conditions for each simulation run.

2) Feature Extraction, Selection, and Classification: To achieve the necessary temporal resolution for real-time UAV swarm control while preserving computational efficiency, we implemented a hybrid signal processing architecture. Specifically, the feature extraction methodology from [16] was selected because its lower dimensionality is essential for mitigating latency within the closed-loop system. For the purpose of continuous neural classification, the incoming EEG streams were partitioned utilizing a 3-second sliding temporal window configured with a 75% overlap.

Inside each respective window, power spectral density values were computed across 10 sub-bands spanning the standard neurological frequency ranges (0.5 Hz $< \delta < 4.0$ Hz, 4 Hz $< \theta < 7.5$ Hz, 8.0 Hz $< \alpha < 13.0$ Hz, 14 Hz $< \beta < 26$ Hz, and 30 Hz $< \gamma < 45.0$ Hz), directly following the conventions detailed in [25]. Consequently, the comprehensive feature vector $\mathbf{X}$ representing a single trial window was formulated as:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1^T & \mathbf{X}_2^T & \cdots & \mathbf{X}_{32}^T \end{bmatrix}^T \quad (1)$$

Here, $\mathbf{X}_\ell \in \mathbb{R}^{10}$ (for $\ell = 1, 2, ..., 32$) denotes the localized feature vector corresponding to each individual electrode among the 32 available channels. Because each channel yields 10 distinct spectral features, the overarching extraction process generates an aggregate of 320 discrete features per second to be utilized by the downstream classifier.

For optimal feature selection, we utilized the Neighborhood Component Analysis (NCA) machine learning algorithm via the MATLAB *fscnca* function, building upon the foundational methodologies established in [16], [17]. A critical methodological advancement in this current study is the implementation of a dynamic, optimal-choice thresholding mechanism for feature retention, which replaces the static threshold approach employed in previous multi-UAV trust evaluations. To determine the most discriminative set of neural features, denoted as $\mathbf{F}$, we systematically evaluated the normalized NCA feature weights ($\mathbf{W}$) against a variable relative threshold factor, r. Specifically, an iterative search was performed where the parameter r was swept from 0.0 to 0.1 in increments of 0.01. During each iteration, a discrete threshold value T was calculated as the product of r and the maximum overall weight. The neural dataset was then temporarily truncated to retain solely the features whose weights exceeded this computed T. Following this reduction, a new NCA model was fitted to the updated validation subset, and its corresponding predictive loss was evaluated against a held-out test set. By continuously tracking these metrics across all candidate values, we successfully identified the optimal relative threshold, r_{best}, that yielded the absolute minimum loss function. This empirically derived threshold ultimately dictated the final feature vector utilized to govern the aerial swarm.

Following the isolation of optimal features, the segmented EEG data was partitioned using a standard 80% training and 20% testing split. To identify the most effective predictive model for our closed-loop architecture, we conducted a rigorous comparative evaluation of three standard supervised learning algorithms: k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and Linear Discriminant Analysis (LDA).

Preliminary offline evaluations demonstrated that the k-NN algorithm yielded superior accuracy and robustness in decoding operator cognitive states; consequently, it was selected as the primary classification engine for the live BCI deployment. All modeling and implementation were executed within the MATLAB environment. To ensure peak algorithmic performance and prevent overfitting, we incorporated automatic hyperparameter optimization coupled with a 5-fold cross-validation protocol, capped at a maximum of 50 objective evaluations, consistent with the robust standards outlined in [17]. Finally, to properly account for the inherent physiological differences in neural oscillations, a completely individualized k-NN classifier was trained exclusively on each participant's data, ensuring highly personalized trust detection during the multi-UAV supervisory tasks.

F. Brain-Computer Interface Phase

Upon the completion of the offline training block, the highly optimized, subject-specific classification model was integrated directly into the real-time BCI architecture for live aerial swarm modulation.

The active BCI framework was engineered to precisely replicate the offline EEGLAB data processing protocols previously detailed, executing identical data windowing and cognitive state classification procedures, as visually summarized in Fig. 4. During active control phases, raw EEG telemetry was initially routed to the local computational unit utilizing the Lab Streaming Layer (LSL) protocol. LSL was implemented continuously throughout our architecture due to its ultra-low latency characteristics (typically below 1 ms), a critical requirement for responsive multi-UAV operations.

This incoming neural data stream was immediately ingested by a dedicated NeuroPype software pipeline. Within this environment, the signals underwent instantaneous preprocessing, encompassing real-time band-pass filtering, dynamic noise suppression via Artifact Subspace Reconstruction (ASR), and common average channel re-referencing to ensure pristine data acquisition. Following the successful completion of these initial filtering steps, the LSL infrastructure was leveraged a second time to transmit the purified data stream directly into a custom Python-based classification module.

Within the Python script, incoming LSL data streams were segmented into 3-second epochs with a 75% overlap. Power spectral density features were extracted and dimensionally reduced using the operator's individualized NCA weights. The k-NN classifier then evaluated this vector to decode the operator's trust level in the UAV swarm. End-to-end system latency averaged 0.25s $\pm$ 0.12s, comfortably below the 0.75-second operational ceiling, ensuring highly responsive real-time control.

III. Results

We present our findings in two sections. First, we detail the offline training phase, analyzing predictive model performance and the neural correlates of multi-UAV trust. Second,

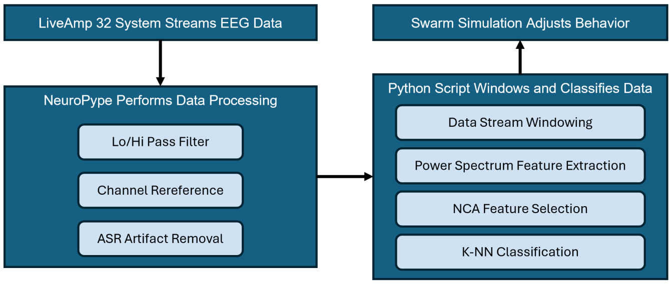

Fig. 4. Architectural overview of the BCI pipeline. Distinct computational blocks represent independent software modules, with their internal subfunctions listed in sequential, top-to-bottom execution order. Directional arrows illustrate the flow of neural data, facilitated entirely by local LSL connections.

we evaluate the active BCI deployment, assessing real-time human-swarm teaming and mission success. For a visual of the training phase and real-time BCI implementation, please see our supplementary video[1].

A. Training

The training phase utilized two distinct stimuli: the goal-oriented *Surround* behavior and the stochastic *Explore* behavior. On average, the *Surround* swarm successfully reached the target within the 25-second limit, whereas the *Explore* behavior consistently failed.

Subjective validation via the Trust Perception Scale [18] corroborated these results; as detailed in Table I, participants rated the *Surround* modality as significantly more trustworthy and reliable than the *Explore* counterpart across all metrics.

TABLE I
TRUST PERCEPTION SCALE IN TRAINING PHASE

Subject	Surround Behavior	Explore Behavior
1	62.25	14.50
2	85.00	19.25
3	51.75	30.50
4	89.25	16.25
5	80.00	27.00

To confirm the presence of neural correlates of trust [16], [17], we utilized a k-NN classifier refined by NCA feature selection. As illustrated in Fig. 5, aggregate classification accuracy exceeded 90%, with the Distrust state yielding slightly higher precision. The NCA protocol optimized the feature space by an average reduction of 55%, retaining 145 out of 320 raw spectral features per participant.

Subsequent analysis of topographical feature maps (Fig. 6) identified the Frontal and Temporal-Parietal regions as the most critical for trust state discrimination. While consistent across most subjects, individual variations occurred, such as the absence of theta and alpha band features in Subject 2's Occipital region. The feature density, mapped on a 2.0 scale, highlights the subject-specific selection weights where red indicates peak feature relevance and blue indicates absence.

[1] https://youtu.be/Lxn6QFppUXM

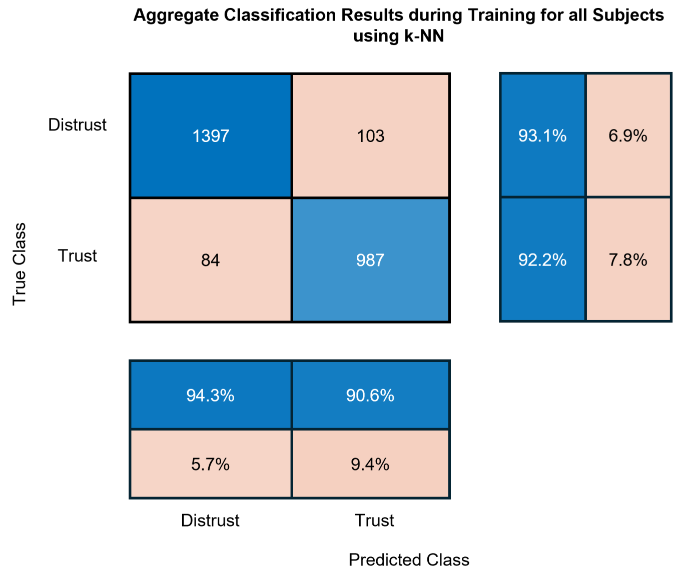

Fig. 5. Aggregate k-NN confusion matrix results across all participants. Utilizing 3-second sliding windows (75% overlap), the experimental trials yielded approximately 500 labeled epochs per subject. Results demonstrate high model robustness, with all participants achieving over 90% prediction accuracy for both *Surround* and *Explore* behavioral states.

B. Real-time BCI

Post-training, participants utilized subject-specific k-NN classifiers to modulate swarm behavior, with performance measured by the duration maintained in the *Surround* state. To minimize false positives, a pre-deployment calibration phase established the number of consecutive "Distrust" windows required to trigger a behavioral shift. As shown in Table II, thresholds varied significantly: Subject 3 required a conservative 6-window buffer, while Subjects 4 and 5 operated optimally with a 1-window trigger.

Figure 7 illustrates control stability via behavior switches per trial. Ideal *Surround* performance (red) is marked by zero or two switches, while successful *Explore* trials (green) typically show an odd switch count, indicating a transition to the trustworthy state. Most subjects maintained stability, though Subjects 4 and 5 showed higher volatility; specifically, Subject 4 registered 20 switches in a single unstable trial, highlighting individual differences in BCI-swarm interfacing.

TABLE II
SUBJECT-SPECIFIC BCI SENSITIVITY THRESHOLDS

Subject	Switching Threshold (Windows)
1	2
2	3
3	6
4	1
5	1

A frequently observed suboptimal pattern is illustrated in Fig. 8 that presents a representative time-series trial. Here, the subject successfully triggered a switch, but the reaction was significantly delayed, occurring nearly halfway through the trial. A similar latency was observed in trials where subjects initially deviated from the desired behavior before recovering,

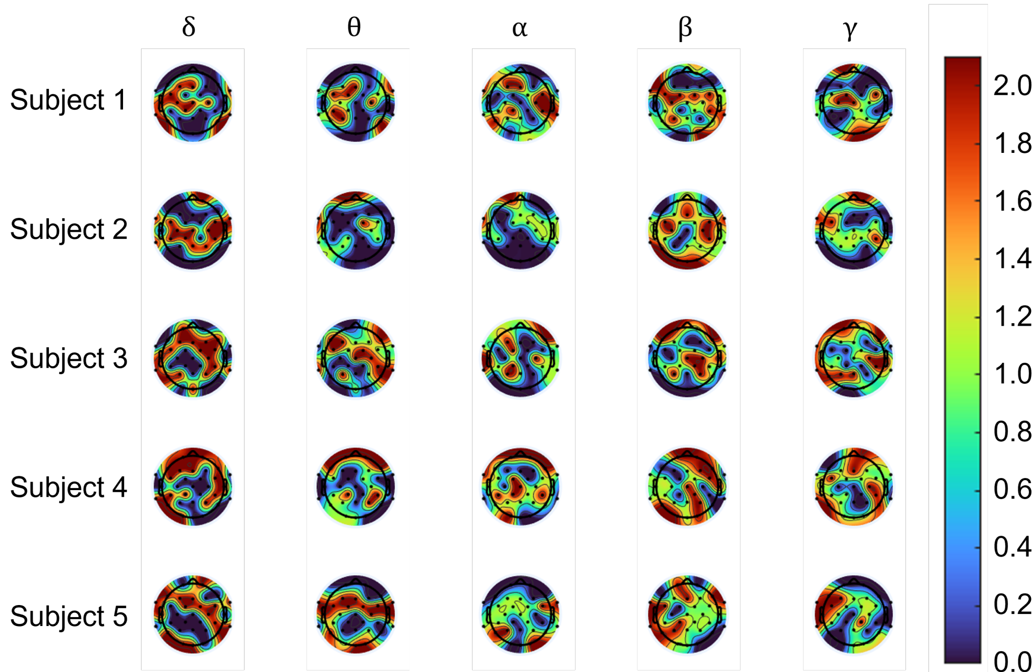

Fig. 6. Topographical distribution of neural feature importance across all subjects (rows) and frequency bands (columns). Scalp map intensity (0–2 scale) indicates the quantity of power-spectral features retained during NCA selection, where dark red (2.0) signifies that both localized sub-band features were identified as critical trust correlates.

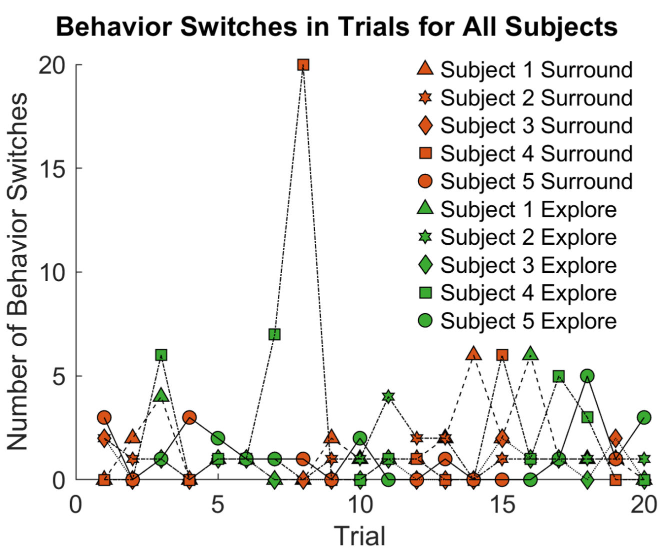

Fig. 7. Quantification of real-time BCI stability via behavior switches per trial. *Surround* trials (red) indicate success through high stability (0 or 2 switches), whereas *Explore* trials (green) demonstrate successful trust transitions via odd switch counts. Most participants exhibited stable control, with the exception of Subject 4, who showed localized instability at Trial 8 with 20 behavior switches.

often losing a third of the trial duration to the wrong state, as shown in Fig. 9.

Conversely, Fig. 10 depicts ideal performance: an immediate, sustained transition to the *Surround* state. This trial also validates our calibration logic; for Subject 2, the 3-window sensitivity threshold successfully prevented a premature behavior switch despite two transient "Distrust" predictions at the 15-second mark, ensuring stable swarm control.

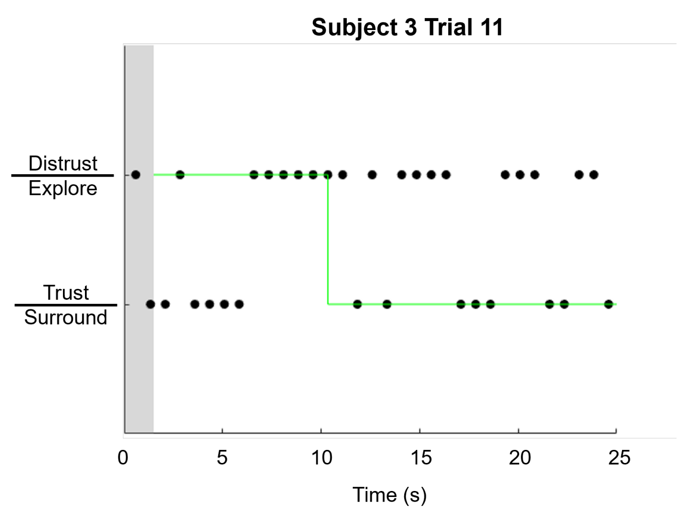

Fig. 8. Representative trial for Subject 3 illustrating delayed control stability. Although the participant achieved the target metric of a single sustained switch, significant transition latency resulted in nearly 40% of the trial duration elapsed before the reliable *Surround* behavior was successfully established.

Following the BCI trials, participants completed the NASA Task Load Index (TLX) [22] to quantify perceived workload on a 0–100 scale. As detailed in Table III, Subject 1 and Subject 2 reported relatively low burdens, while Subject 4 perceived moderate demand. High workload levels were recorded for Subject 5 and Subject 3. Notably, Subject 3 reported the highest overall index, characterized by extreme Mental Demand (95), Effort (90), and Frustration (85).

IV. Conclusion

This study introduces a novel EEG-based BCI framework that operationalizes human trust for closed-loop human-swarm

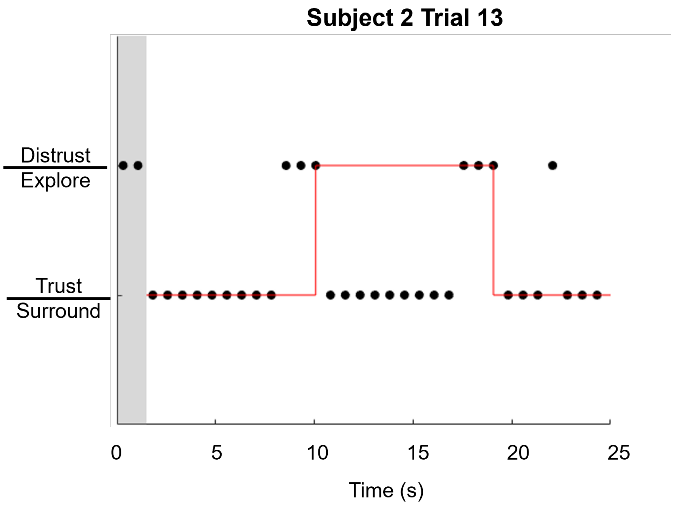

Fig. 9. Real-time performance for Subject 2 demonstrating recovery of the trustworthy state. Despite an initial deviation that limited the *Surround* behavior to 60% of the trial, the participant successfully re-established and sustained the target behavior for the remainder of the session.

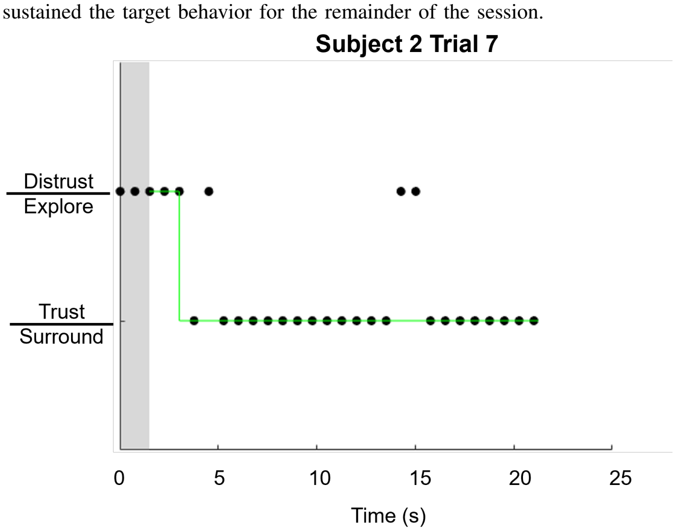

Fig. 10. Representative trial for Subject 2 demonstrating near-ideal BCI control. The participant executed a single, rapid transition at the earliest possible window, maintaining the trustworthy *Surround* behavior for approximately 90% of the trial duration.

TABLE III
NASA TLX RATING SCALE RESULTS

Item	**S1**	**S2**	**S3**	**S4**	**S5**
Mental Demand	60.00	60.00	95.00	70.00	95.00
Physical Demand	05.00	05.00	05.00	30.00	10.00
Temporal Demand	35.00	45.00	70.00	25.00	75.00
Performance	25.00	15.00	60.00	55.00	70.00
Effort	30.00	55.00	90.00	85.00	80.00
Frustration	25.00	15.00	85.00	30.00	70.00
Total	**34.33**	**39.33**	**83.33**	**62.67**	**73.67**

interaction, bridging a critical gap in intuitive interface design. Utilizing a k-NN classifier, the system enables a virtual swarm to perceive operator distrust and dynamically adapt behavior, fostering a symbiotic partnership. While real-time trials demonstrated successful control, significant performance variability revealed key limitations. A primary challenge was the discrepancy between the training phase, which focused on maintaining static mental states, and the real-time task, which required dynamic executive transitions. This suggests the classifier was optimized for stability rather than state-switching, likely contributing to the performance bias observed between maintenance and transition trials.

Furthermore, the difficulty of verifying "ground truth" for passive, continuous cognitive states, unlike discrete motor or speech commands, introduces label noise from attentional drift, exacerbated by high mental workload. Future work will prioritize dynamic calibration protocols that capture neural features specific to initiating behavioral changes. Additionally, we propose investigating hybrid architectures that incorporate event-locked signals, such as Error-Related Potentials (ERPs) [26], [27], to provide distinct neural signatures for perceived errors, thereby reducing cognitive load and improving classification robustness.

Ultimately, this research establishes a vital proof-of-concept for high-level affective BCIs in aerial robotics. By shifting from low-level motor commands to abstract cognitive states, we enable a more natural, hands-free collaboration between humans and autonomous swarms. This work lays the foundation for a new generation of interfaces where robotic swarms dynamically adapt to the operator's internal state, ensuring seamless and resilient teaming in high-stakes domains such as search and rescue or defense.

REFERENCES

[1] S. Park, "Smarter robot swarms could change warfare," Aug 2021. [Online]. Available: https://www.ausa.org/articles/smarter-robot-swarms-could-change-warfare

[2] V. Krátkỳ, G. Silano, M. Vrba, C. Papaioannidis, I. Mademlis, R. Pěnička, I. Pitas, and M. Saska, "Gesture-controlled aerial robot formation for human-swarm interaction in safety monitoring applications," *IEEE robotics and automation letters*, 2025.

[3] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," in *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, 1987, pp. 25–34.

[4] A. J. Hepworth, D. P. Baxter, A. Hussein, K. J. Yaxley, E. Debie, and H. A. Abbass, "Human-swarm-teaming transparency and trust architecture," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 7, pp. 1281–1295, 2020.

[5] K. A. Roundtree, J. R. Cody, J. Leaf, H. O. Demirel, and J. A. Adams, "Transparency's influence on human-collective interactions," *J. Hum.-Robot Interact.*, vol. 11, no. 2, Mar. 2022. [Online]. Available: https://doi.org/10.1145/3507470

[6] E. S. Kox, J. van den Boogaard, V. Turjaka, and J. H. Kerstholt, "The journey or the destination: The impact of transparency and goal attainment on trust in human-robot teams," *J. Hum.-Robot Interact.*, vol. 14, no. 2, Dec. 2024. [Online]. Available: https://doi.org/10.1145/3702245

[7] G. K. Karavas, D. T. Larsson, and P. Artemiadis, "A hybrid bmi for control of robotic swarms: Preliminary results," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sep. 2017, pp. 5065–5075.

[8] C. H. Nguyen, G. K. Karavas, and P. Artemiadis, "Inferring imagined speech using eeg signals: a new approach using riemannian manifold features," *Journal of Neural Engineering*, vol. 15, no. 1, p. 016002, dec 2017. [Online]. Available: https://doi.org/10.1088/1741-2552/aa8235

[9] ——, "Adaptive multi-degree of freedom brain computer interface using online feedback: Towards novel methods and metrics of mutual adaptation between humans and machines for bci," *PLOS ONE*, vol. 14, no. 3, pp. 1–25, 03 2019. [Online]. Available: https://doi.org/10.1371/journal.pone.0212620

[10] J. D. Lee and K. A. See, "Trust in automation: Designing for appropriate reliance," *Human factors*, vol. 46, no. 1, pp. 50–80, 2004.

[11] W.-L. Hu, K. Akash, T. Reid, and N. Jain, "Computational modeling of the dynamics of human trust during human–machine interactions," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 6, pp. 485–497, 2019.

[12] F. Ekman, M. Johansson, and J. Sochor, "Creating appropriate trust in automated vehicle systems: A framework for hmi design," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 1, pp. 95–101, 2018.

[13] S. Choo and C. S. Nam, "Detecting human trust calibration in automation: A convolutional neural network approach," *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 4, pp. 774–783, 2022.

[14] Y. Guo, Y. Pang, J. Lyons, M. Lewis, K. Sycara, and R. Liu, "Trust-aware reflective control for fault-resilient dynamic task response in human–swarm cooperation," *AI*, vol. 5, no. 1, pp. 446–464, 2024.

[15] C. Nam, P. Walker, H. Li, M. Lewis, and K. Sycara, "Models of trust in human control of swarms with varied levels of autonomy," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 3, pp. 194–204, 2020.

[16] J. A. Orozco and P. Artemiadis, "Extracting human levels of trust in human–swarm interaction using EEG signals," *IEEE Transactions on Human-Machine Systems*, vol. 54, no. 2, pp. 182–191, 2024.

[17] J. A. Orozco, S. Bachoumas, and P. Artemiadis, "Quantifying human trust in controlling robot swarms: Eeg-based analysis and classification," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, vol. 2025, 2025, pp. 1–7.

[18] K. E. Schaefer, "Measuring trust in human robot interactions: Development of the "trust perception scale-hri"," in *Robust intelligence and trust in autonomous systems*. Springer, 2016, pp. 191–218.

[19] G. H. Klem, "The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology," *Electroencephalogr. Clin. Neurophysiol. Suppl.*, vol. 52, pp. 3–6, 1999.

[20] Unity 20223.3.17f1, Unity Technologies, 2025. [Online]. Available: https://unity.com/

[21] C. Kothe, "Lab streaming layer," 2013. [Online]. Available: https://labstreaminglayer.readthedocs.io/

[22] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*. Elsevier, 1988, vol. 52, pp. 139–183.

[23] D. Arnaud and M. Scott, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, pp. 9–21, 2004. [Online]. Available: http://www.sccn.ucsd.edu/eeglab/

[24] T. R. Mullen, C. A. Kothe, Y. M. Chi, A. Ojeda, T. Kerth, S. Makeig, T.-P. Jung, and G. Cauwenberghs, "Real-time neuroimaging and cognitive monitoring using wearable dry eeg," *IEEE transactions on biomedical engineering*, vol. 62, no. 11, pp. 2553–2567, 2015.

[25] S. Sanei and J. A. Chambers, *EEG signal processing and machine learning*. John Wiley & Sons, 2021.

[26] P. W. Ferrez and J. d. R. Millán, "Error-related eeg potentials generated during simulated brain–computer interaction," *IEEE transactions on biomedical engineering*, vol. 55, no. 3, pp. 923–929, 2008.

[27] A. F. Salazar-Gomez, J. DelPreto, S. Gil, F. H. Guenther, and D. Rus, "Correcting robot mistakes in real time using eeg signals," in *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2017, pp. 6570–6577.