Predictive Modeling of G-Induced Loss of Consciousness (GLOC) Using Physiological Data and Machine Learning

Lucas R. Owen, Joshua B. Smith, Eunjun Shin, and Sarah Bolton
Department of Computer and Cyber Sciences
United States Air Force Academy
Colorado, USA

***Abstract*—G-induced Loss of Consciousness (GLOC) is a persistent hazard in high-performance aviation that can result in catastrophic loss of control. This study develops a machine learning model using behavioral and physiological signals to predict imminent GLOC from centrifuge trial data provided by the Air Force Research Laboratory (AFRL). After systematically eliminating sources of data leakage, including centrifuge telemetry, device accelerometers, motion-contaminated EEG signals, and an experimental condition confound, a leakage-free Random Forest classifier trained exclusively on Non-AFE rapid-onset runs achieves a Recall of 0.86, Precision of 0.72, and F1 score of 0.78 in leave-one-subject-out cross-validation. Key predictive signals include cerebral oxygenation (fNIRS), joystick tracking degradation, and pupillometry. Results demonstrate that cockpit-deployable sensors alone are sufficient to predict GLOC onset with clinically meaningful accuracy.**

***Index Terms*—GLOC, machine learning, physiological sensing, fNIRS, pupillometry, aviation safety, Random Forest**

I. Introduction

High-performance aviation imposes extreme physiological and cognitive demands on aircrew during high-G maneuvers. G-induced Loss of Consciousness (GLOC) occurs when sustained acceleration forces reduce cerebral blood flow, causing temporary incapacitation. Even transient GLOC episodes can lead to catastrophic events such as Controlled Flight Into Terrain (CFIT). While Auto-GCAS has saved numerous lives, its deterministic logic operates reactively, intervening only after ground impact becomes imminent. To further enhance pilot survivability, predictive systems capable of assessing pilot physiological states in real-time are required.

This research integrates physiological sensor data and machine learning to predict cognitive degradation under G-stress conditions. By modeling behavioral and physiological time-series signals (joystick tracking, pupillometry, and cerebral oxygenation), we predict GLOC onset before complete incapacitation occurs. A central methodological contribution is the systematic identification and elimination of data leakage sources, producing a defensible and generalizable model.

II. Background

A. Pilot Disorientation and GLOC

GLOC remains a major operational hazard in tactical aviation. Even brief unconsciousness at high speed can result in irreversible loss of control. Auto-GCAS, introduced in 2014, has saved 13 pilots and 12 aircraft to date [1], [2]. However, its deterministic algorithms are not designed to adapt to pilot physiological state. In low-altitude or evasive flight profiles, false alarms and early activation remain challenges [3]. The next generation of flight safety systems must integrate probabilistic models capable of predicting physiological incapacitation and initiating context-sensitive interventions.

B. Physiological Sensing and Cognitive Assessment

Modern wearable biosensors capture diverse physiological and cognitive parameters in real time. Electroencephalography (EEG) and electrooculography (EOG) provide insight into pilot workload and visual engagement [4], [5]. ECG, photoplethysmography (PPG), and electrodermal activity (EDA) signals estimate stress, fatigue, and cognitive load [6], [7]. Recent studies demonstrated that convolutional neural networks (CNNs) trained on multimodal physiological signals can classify emotional and cognitive states with accuracies exceeding 75% [8]. Within aerospace contexts, the AFRL's Integrated Cockpit System (ICS) has facilitated controlled centrifuge trials capturing EEG, heart rate, breathing rate, and gaze data from aircrew [9].

C. Machine Learning for Predicting GLOC

Several recent efforts have applied machine learning to G-tolerance and GLOC. Rinkel [10] reported approximately 85% accuracy predicting GLOC onset up to 15 seconds in advance using EEG, ECG, and respiration. Ohrui et al. [12] trained support vector machines on age, height, weight, anti-G suit status, $+G_z$ level, and cerebral oxy- and deoxyhemoglobin, predicting GLOC with roughly 65% accuracy from 2.5 seconds after the onset of high-G_z exposure. G-tolerance has also been predicted from mobile heart rate and activity data alone [11].

The present work differs from these studies in two ways that matter for interpretation. First, the prediction target is different. Ohrui et al. predict *whether* a given subject will experience GLOC, a subject-level binary classification on balanced classes evaluated with leave-one-out cross-validation, and report roughly 65% accuracy, which the authors note is not yet sufficient for operational use. Our task instead predicts *when*: each timestep is classified as pre-GLOC or not within a population of subjects who all experience GLOC, so we report Recall, Precision, and F1 on an imbalanced within-run window rather than balanced accuracy, and a direct numerical comparison is not meaningful. Both evaluations are subject-independent; ours uses leave-one-subject-out cross-validation, which tests every model on a pilot it has never seen. Second, our leakage-controlled model attains strong recall using only joystick tracking and pupillometry, signals compatible with an operational cockpit, whereas the strongest prior GLOC-prediction accuracies rely on less deployable instrumentation such as the EEG, ECG, and respiration channels used by Rinkel [10]. Notably, cerebral oxygenation emerges as a leading predictor in both Ohrui's analysis and ours, and Ohrui explicitly identify Random Forests as a promising future direction, the approach adopted here. Translating any of these findings into deployable systems requires models robust to sensor noise, limited data availability, and the physical constraints of pilot gear [13]. The present research uses AFRL's centrifuge dataset, containing time-synchronized physiological and behavioral data, to train and evaluate ensemble classifiers.

D. Toward Cognitive-Aware Flight Safety Systems

Interviews with AFRL and DoD subject-matter experts highlight a persistent research gap: while current efforts focus on detecting GLOC occurrence, limited work addresses predicting recovery or post-GLOC cognitive capability. A 'cognitive safety layer' could enable autonomous systems to adaptively assist or assume control based on pilot state estimates, providing an anticipatory safeguard that complements Auto-GCAS.

III. Dataset

The dataset was collected by the AFRL 711th Human Performance Wing during controlled human centrifuge trials [9]. It contains synchronized multimodal recordings from subjects undergoing gradual onset rate (GOR) and rapid onset rate (ROR) $+G_z$ profiles, approximately 1.4 million rows across 61 subjects and 312 columns. Modalities included:

- **Centrifuge telemetry:** G-force profiles, G-suit pressure, oxygen delivery
- **EEG** (Brain Products 32-lead, 500 Hz): Raw signals and pre-computed delta, theta, alpha, and beta band power
- **Eye tracking** (Tobii glasses): Pupil diameter, pupil position in 3D space, IMU accelerometers
- **Joystick / PTracker:** Position, velocity, deviation from target
- **fNIRS:** Cerebral deoxyhemoglobin (Hbd) and oxyhemoglobin (HbO2)
- **Physiological** (Equivital LifeMonitor): Heart rate, breathing rate, skin temperature, accelerometers
- **Math task:** Response time, accuracy, problem type

Two experimental conditions were present: **Non-AFE** (no anti-G augmentation; 34/34 ROR runs resulted in GLOC) and **AFE** (anti-G suit with augmented flow equipment; only 5/27 ROR runs resulted in GLOC).

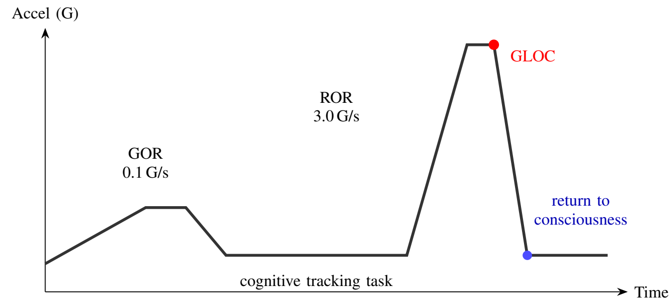

Fig. 1. Centrifuge acceleration profile per subject: a gradual onset run establishes baseline tolerance, then a rapid onset run at $3.0\,G/s$ drives acceleration to GLOC while the subject performs a cognitive tracking task and the AGSM.

During each run, subjects performed a continuous joystick pursuit-tracking task and an intermittent mental-math task, providing the behavioral (tracking degradation) and cognitive (response time, accuracy) channels used below. Each subject followed the same protocol (Fig. 1): after a cognitive-task refresher and EEG baseline, a gradual onset run (GOR) at approximately $0.1\,G/s$ established a baseline G tolerance, followed by lower-body strain engagement; a rapid onset run (ROR) at approximately $3.0\,G/s$ then raised acceleration until GLOC while the subject performed the cognitive tracking task and the anti-G straining maneuver (AGSM), with a post-GLOC cognitive task after return to consciousness. All subjects therefore experienced the same profile and task load, though detailed per-subject demographics (age, sex, and flight experience) were not included in the dataset documentation available to the authors. The raw data is AFRL-internal; the same 711th Human Performance Wing centrifuge collection is described in the publicly available thesis of Rinkel [10], which readers may consult for further procedural detail.

IV. Methodology

A. Problem Framing Evolution

The initial goal was direct GLOC onset prediction. The team subsequently pivoted to predicting post-GLOC recovery, exploring holistic physiological approaches, EEG combined with task data, and tasks with eye tracking, each abandoned due to between-subject variability, EEG noise, or data sparsity. A critical finding during this phase was that the math task had a 95% correct-response rate, making binary accuracy an unusable label; response time emerged as a more sensitive continuous measure. After determining the dataset did not adequately support recovery prediction, the team returned to GLOC onset prediction with a rigorous, leakage-free methodology.

B. Data Leakage Investigation

A central methodological theme was systematically identifying and removing features that reflected the G-force environment rather than the pilot's genuine physiological state.

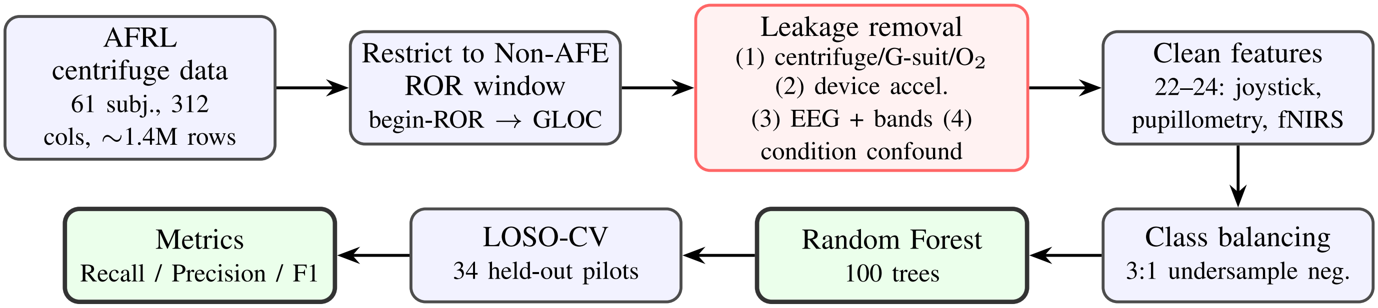

Fig. 2. End-to-end machine learning pipeline. Raw multimodal centrifuge data is restricted to the Non-AFE rapid-onset window, four sources of data leakage are removed (red), and the resulting 22 to 24 clean features feed a class-balanced Random Forest evaluated with leave-one-subject-out cross-validation.

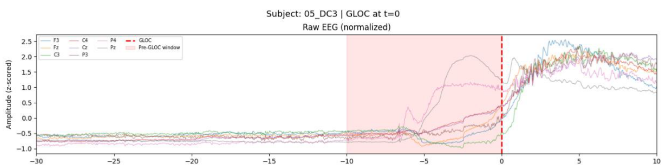

Fig. 3. EEG motion artifact at GLOC onset across parietal electrodes.

1) Centrifuge and G-Force Columns: Initial models using all 213 features were dominated by centrifuge telemetry columns encoding the programmed G profile of the run, not the pilot's response to it. A model learning these features is essentially reading the centrifuge script rather than monitoring the pilot. All centrifuge, G-suit, and oxygen delivery columns were excluded.

2) Device Accelerometers: After removing centrifuge columns, the top features shifted to EEG headset and Tobii IMU accelerometer outputs, physical G-force measurements on the devices, not brain or eye signals. All device accelerometer columns were excluded.

3) EEG Motion Artifacts: EEG delta band power (particularly parietal electrodes P3 to P8 and Pz) emerged next. Visual inspection across all 39 GLOC events revealed that delta power was flat throughout the pre-GLOC window and spiked at or after the GLOC timestamp, a textbook motion artifact caused by involuntary head movement at loss of muscle tone. All EEG channels and band power features were excluded.

4) Condition Confound: GLOC occurrence was almost perfectly predicted by condition alone (34/34 Non-AFE vs. 5/27 AFE). Any model trained across both conditions primarily learns the condition label rather than pilot-state signals. Analysis was restricted to Non-AFE runs only, where all pilots experience GLOC and the prediction task becomes *when*, not *whether*.

5) Prediction Window: Pilots typically experience GLOC within 7 to 20 seconds of beginning a rapid onset run (median: 12 seconds). Data was filtered to only include rows between the begin-ROR and GLOC events, ensuring all training data reflects genuine high-G pilot state.

C. Final Feature Set

After eliminating all identified sources of leakage, the final clean feature set consisted of 22 to 24 features across three modalities:

- **Joystick / PTracker:** joystickX, joystickY, joystickPosMag, joystickVelMag, deviation, userdeltaX, userdeltaY, targetDeltaX, targetDeltaY, tgtposX, tgtposY, timeDelta, mouseD1, mouseD2
- **Pupillometry** (eye tracking, IMU excluded): Pupil diameter left/right; Pupil position left/right X/Y/Z
- **fNIRS** (when available): Hbd (deoxyhemoglobin), HbO2 (oxyhemoglobin)

fNIRS coverage showed a clean binary split: 22 of 34 Non-AFE subjects had 100% coverage within the ROR window; 12 had 0% coverage, consistent with sensor disconnection noted in the data dictionary.

D. Model Architecture and Evaluation

The complete pipeline is summarized in Fig. 2. A **Random Forest Classifier** (100 trees) was selected for robustness out-of-the-box, interpretability via feature importances, and tolerance to mixed feature types. The prediction task is binary classification: does this timestep fall within 10 seconds before a GLOC event?

Class balancing addressed severe imbalance (8,684 positive vs. 4,144 negative samples within ROR windows) via 3:1 undersampling of negatives during training.

Implementation. Models were implemented in Python with scikit-learn, using `RandomForestClassifier` (100 trees) for the final model and `LeaveOneGroupOut` for cross-validation; the gradient-boosted comparison used XGBoost with GPU acceleration. A trained Random Forest of this size performs inference at low computational cost, consistent with the demands of real-time cockpit operation.

Evaluation protocol. Performance was measured with **leave-one-subject-out cross-validation (LOSO-CV):** the model was trained on all subjects except one, tested on the held-out subject, and repeated for every subject, the only evaluation method that honestly tests generalization to unseen pilots. We report Recall (fraction of true pre-GLOC timesteps recovered), Precision (fraction of positive predictions that were correct), and F1 (their harmonic mean); accuracy is not reported because the class imbalance within ROR windows makes it uninformative.

Alternative architectures explored included XGBoost with GPU acceleration (comparable performance, more tuning required) and temporal feature engineering via rolling statistics over 5 s, 10 s, and 20 s windows (no improvement; added noise and complexity).

E. Dead Ends and Alternative Modeling Approaches

A substantial portion of this research consisted of exploring modeling directions that ultimately proved ineffective but were critical in shaping the final methodology. These "dead ends" revealed key structural limitations of the dataset and informed the design of a leakage-free, generalizable model.

1) Sequential Deep Learning Approaches (LSTM / RNN): Initial efforts focused on recurrent neural networks (RNNs), including Long Short-Term Memory (LSTM) architectures, due to the inherently temporal nature of physiological signals. Early results appeared highly promising, achieving near-perfect training and validation accuracy. However, further investigation revealed that these results were driven by severe data leakage and improper windowing. Specifically, overlapping sliding windows caused identical temporal segments from the same trial to appear in both training and validation sets, artificially inflating performance. After correcting this by enforcing strict trial-level separation, model performance collapsed due to extreme data sparsity and class imbalance. In some configurations, fewer than 20 valid training windows remained, resulting in unstable training (NaN loss) and no meaningful generalization.

Additionally, the dataset exhibited a structural imbalance in cognitive task labels: approximately 92% of responses were correct, rendering accuracy an unreliable metric. Models trained on this label distribution converged to trivial majority-class predictions, highlighting that classification of task correctness is not a viable proxy for cognitive degradation.

2) Cognitive Task-Based Modeling: A parallel line of investigation attempted to predict GLOC or cognitive impairment using math task performance (accuracy and response behavior). This approach was ultimately abandoned due to three key limitations:

1) **Ceiling effect:** near-perfect accuracy eliminated discriminative signal,
2) **Data sparsity:** only $\sim$31% of timesteps contained task data, and
3) **Temporal misalignment:** task responses were not consistently synchronized with physiological degradation.

While classification proved ineffective, this exploration revealed that response time is a more sensitive indicator of cognitive impairment than binary correctness. This insight remains a promising direction for future work on post-GLOC recovery modeling.

3) Feature Engineering with Physiological Signals (HR, BR, Temperature): Early models incorporated traditional physiological signals such as heart rate, breathing rate, and skin temperature. Despite their theoretical relevance to stress and G-tolerance, these features contributed minimal predictive power in practice. This aligns with SME feedback indicating that such signals are either too slow-changing or too noisy under high-G conditions to capture rapid cognitive degradation. As a result, these features were deprioritized in favor of higher-frequency behavioral and neurovascular signals.

4) Boosting Methods and Alternative Classical Models: Several alternative models were explored, including XGBoost, gradient boosting, k-nearest neighbors (KNN), and support vector machines (SVM). While boosting methods achieved performance comparable to Random Forests, they required extensive hyperparameter tuning and exhibited greater sensitivity to noise and class imbalance. KNN and SVM approaches struggled to scale to the high-dimensional, time-series nature of the dataset and failed to capture nonlinear temporal dependencies effectively. Ultimately, Random Forests provided the best balance of robustness, interpretability, and performance with minimal tuning.

5) Temporal Feature Engineering: To better capture time-dependent dynamics, rolling statistical features (mean, variance, slope) were computed over 5 s, 10 s, and 20 s windows. Contrary to expectations, these features degraded model performance. The likely cause is that temporal aggregation smoothed out short-term physiological changes immediately preceding GLOC, which are critical for prediction. This suggests that instantaneous deviations, rather than long-window trends, are the most informative signals in rapid-onset G-force scenarios.

6) Unsupervised and Clustering Approaches: Exploratory work considered unsupervised learning techniques such as clustering to identify latent physiological states. However, high inter-subject variability and inconsistent sensor coverage prevented the emergence of stable, interpretable clusters. This reinforces the necessity of supervised approaches with careful subject-level validation.

V. Results

Table I summarizes all model runs in chronological order of development. Each row represents a substantive methodological change. The highlighted row (*) indicates the best-performing configuration.

A. Event-Level Detection

Table I reports timestep-level metrics, which is the level at which the model was trained and evaluated. Because an operational system alarms per *event* rather than per timestep, the natural complementary metric is event-level detection: for each of the 34 Non-AFE GLOC events, whether at least one pre-GLOC alarm is raised inside the warning window, together with the associated false-alarm rate per run and the warning lead time. This event-level characterization, computed by aggregating the timestep predictions under a persistence rule that suppresses isolated positives, is the immediate next step in this analysis and is discussed qualitatively in Section VII.

TABLE I

MODEL PERFORMANCE ACROSS METHODOLOGY ITERATIONS. RECALL = FRACTION OF PRE-GLOC EVENTS CORRECTLY IDENTIFIED; PRECISION = FRACTION OF PRE-GLOC PREDICTIONS THAT WERE CORRECT; F1 = HARMONIC MEAN.

Configuration	Recall	Prec.	F1
Handpicked features (30)	0.82	0.46	0.59
All 213 features	0.63	0.50	0.56
No centrifuge columns (198)	0.66	0.33	0.44
Pure pilot-state (168)	0.62	0.21	0.31
Temporal XGBoost (540)	0.52	0.26	0.35
Temporal RF 20 s (162)	0.45	0.29	0.35
No EEG / no accelerometers (24)	0.39	0.16	0.22
Non-AFE ROR only, no fNIRS (22)	0.86	0.67	0.75
Non-AFE ROR only, with fNIRS (24)*	**0.87**	**0.72**	**0.78**
fNIRS subjects only (24)	0.80	0.68	0.73

VI. KEY FINDINGS

A. *fNIRS Is the Strongest Physiological Predictor*

When restricted to the clean Non-AFE ROR window, Hbd and HbO2 were the top two features, accounting for approximately 23% of feature importance. This is physiologically meaningful: fNIRS directly measures cerebral blood oxygenation, which is the biological mechanism of GLOC. As G-forces build, cerebral blood flow drops, hemoglobin delivery falls, and the brain becomes hypoxic.

B. *The Model Works Without fNIRS*

Removing fNIRS reduced recall only marginally (0.87 to 0.86) while precision dropped from 0.72 to 0.67. Joystick deviation and pupil tracking alone are sufficient for strong performance; both sensors could realistically be deployed in an actual cockpit without specialized neuroimaging equipment. This is operationally significant.

C. *EEG Is Not Usable in This Environment*

Visual inspection of EEG delta band power around all 39 GLOC events confirmed motion artifact contamination. Delta power was flat during the pre-GLOC window and spiked at or after GLOC onset across the majority of subjects. EEG cannot serve as a predictive feature without artifact rejection preprocessing specifically tailored to the G-force environment.

D. *Condition Determines GLOC More Than Any Physiological Signal*

The AFE/Non-AFE condition is the single strongest predictor of GLOC in this dataset. Any model that does not account for this confound will learn condition membership rather than pilot state. Restricting to Non-AFE runs is essential for honest evaluation.

E. *Final Performance*

Using only joystick and pupillometry features, within Non-AFE ROR windows, evaluated with leave-one-subject-out cross-validation across 34 pilots:

- **Recall: 0.86** the model correctly identifies 86% of pre-GLOC timesteps
- **Precision: 0.67** 67% of pre-GLOC predictions are correct
- **F1: 0.75**

Including fNIRS for the 22 subjects with sensor coverage improves performance to Precision 0.72, F1 0.78. This represents a defensible, leakage-free result demonstrating that joystick tracking degradation and pupil behavior changes are genuine predictive signals for GLOC onset in rapid onset G-force runs.

VII. OPERATIONAL DEPLOYMENT IMPLICATIONS

A. *Warning Timeline*

The prediction target is a 10-second pre-GLOC window, and in this dataset GLOC follows rapid onset within 7 to 20 seconds (median 12). Because the model raises positive predictions inside this pre-onset interval, the achievable warning lead time is on the order of a few seconds, which is enough to matter operationally. Roughly three seconds is sufficient to prompt or reinforce the anti-G straining maneuver, the interval at which the maneuver is normally cycled, while even one to two seconds of anticipation is enough to prime an automated recovery handoff. Shortening the lookahead window from 10 to 5 seconds is expected to raise precision (positives sit closer to genuine onset) at the cost of lead time; because a median run reaches GLOC in 12 seconds, a 5-second window still leaves a usable margin and warrants systematic evaluation. The window length should therefore be treated as a tunable operating point rather than a fixed choice, and quantifying the resulting event-level lead-time distribution is a direct next step.

B. *Sensor Deployability*

The two features that carry most of the non-fNIRS signal, joystick tracking degradation and pupillometry, map directly onto hardware compatible with an operational cockpit. Joystick inputs are already instrumented in fly-by-wire aircraft, and pupil metrics are obtainable from helmet-mounted or panel eye-tracking. This aligns with AFRL subject-matter-expert feedback that eye tracking was among the most informative and least invasive available signals and the most plausible to field. fNIRS improves precision (0.67 to 0.72) but is not required for strong recall, so a first-generation system need not carry neuroimaging hardware. This lowers the integration barrier substantially relative to prior approaches built on dense EEG or chest-mounted physiological arrays.

C. *Integration with Existing Safety Systems*

These results support a probabilistic cognitive safety layer that complements, rather than replaces, deterministic protection. On a rising GLOC probability the layer could escalate through graded responses: prompt or reinforce the anti-G straining maneuver, pre-arm Auto-GCAS so recovery logic engages with less delay, and, on high-confidence predicted incapacitation, hand control to onboard autonomy until the pilot recovers. The cost asymmetry between a nuisance alarm and an undetected GLOC favors tuning toward recall, with a short persistence or confirmation gate to hold false alarms to an acceptable per-run rate. Establishing that operating point on a larger and more demographically varied subject pool is the primary prerequisite for flight-representative evaluation.

VIII. Limitations and Future Work

- **Small positive class:** Only 39 GLOC events across 61 subjects limits model generalizability. Results should be interpreted with caution.
- **fNIRS reliability:** 35% of subjects had no fNIRS data due to sensor disconnection. A more robust fNIRS setup could significantly improve performance.
- **EEG unused:** With proper artifact rejection preprocessing tailored to the G-force environment, EEG band power could be a valuable signal. The parietal delta increase seen post-GLOC suggests genuine neural signal may be present but obscured by motion artifacts.
- **Lookahead window:** The 10-second lookahead was not systematically optimized. Given a median ROR duration of 12 seconds, earlier warning windows (5 seconds) should be explored.
- **Subject variability:** LOSO-CV likely underestimates true generalization error given the small subject count and high between-subject physiological variability.
- **Response time not explored:** The finding that response time is a more sensitive cognitive measure than task accuracy was not fully pursued. This remains a promising direction for post-GLOC recovery prediction.

IX. Conclusion

This study demonstrates that GLOC onset can be predicted with strong performance using only cockpit-deployable sensors (joystick tracking and pupillometry) when a rigorous methodology eliminates data leakage. The systematic identification of centrifuge telemetry proxies, device accelerometer contamination, EEG motion artifacts, and an experimental condition confound was essential: models naively trained on all available features vastly overestimated performance.

Restricting analysis to Non-AFE rapid-onset runs and evaluating with leave-one-subject-out cross-validation yields a final Recall of 0.86, Precision of 0.67 to 0.72, and F1 of 0.75 to 0.78 depending on fNIRS availability. The convergence of physiological sensing, machine learning, and flight control presents a transformative opportunity for aerospace safety. By extending deterministic safety systems with probabilistic, data-driven pilot state prediction, this research enables intelligent autonomy that cooperatively preserves both pilot and mission integrity.

References

[1] A. Church, "The Science of Avoidance," *Air & Space Forces Magazine*, 2016.

[2] "Automatic Ground Collision Avoidance System," Air Force Research Laboratory, 2019.

[3] S. Lee, "Modified Algorithm of Predictive Ground Collision Avoidance System for the Low-Level Terrain Flight," IEEE, 2010.

[4] C. Belkhiria and V. Peysakhovich, "EEG and EOG in Aeronautics: A Review," *Frontiers in Neuroergonomics*, 2020.

[5] W. Romie et al., "Measuring Cognitive Load Using Electrodermal and Physiological Sensors," *Sensors*, 2020.

[6] C.-J. Yang et al., "An AI-Edge Platform with Multimodal Physiological Signals," IEEE, 2020.

[7] S. Kim et al., "Wearable Biosignal Monitoring System Applied to Aviation Safety," IEEE, 2017.

[8] P. Moore et al., "Efficient Deep Learning on Wearable Sensor Data for Pilot Flight Performance," *IEEE BSN*, 2023.

[9] AFRL, "Integrated Cockpit System Data Collection," Unpublished internal documentation, 2024.

[10] B. Rinkel, "G-Induced Loss of Consciousness Predictive Model Development," WPI, 2024.

[11] M.-H. Kuo, "G Tolerance Prediction Using Mobile-Device Measured Cardiac Force Index," *JMIR MHealth UHealth*, 2023.

[12] N. Ohrui, Y. Iino, K. Kuramoto, A. Kikukawa, K. Okano, K. Takada, and T. Tsujimoto, "G-Induced Loss of Consciousness Prediction Using a Support Vector Machine," *Aerospace Medicine and Human Performance*, vol. 95, no. 1, pp. 29–36, 2024.

[13] D. Shaw and J. Harrell, "Integrating Physiological Monitoring Systems in Military Aviation," *Ergonomics*, 2023.