

A method for analyzing biometric data to assess the cognitive load and stress resistance of an agricultural unmanned aerial vehicle operator during mission planning

L.A. Taskina¹, L.A. Abakumov¹, T.D. Kazarkin¹

¹ Samara National Research University, Moskovskoye Shosse 34, Samara, 443086, Russia

Abstract

We present a computer vision method to assess cognitive load and stress resistance of an agricultural unmanned aerial vehicle operator during mission planning in virtual reality. The approach combines geometrically rigorous gaze-to-user-interface mapping projecting the eye-tracker ray into widget space to obtain metrically correct hits on areas of interest, behavioral and ocular biomarkers, and image-like attention representations, such as heatmaps and recurrence plots. In a study with twelve participants across four scenarios, we recorded 1,198 interaction events and obtained 85.3% accuracy of gaze-to-interface hits; with increasing difficulty, fixation durations shortened, transition entropy increased, and event-locked pupil responses became larger and slower to recover. Planning time and the time required for replanning increased, while route quality decreased under time pressure. The approach relies only on software-platform aggregate signals and does not use raw eye images, which supports privacy-preserving deployment and portability to ground control software.

Keywords: agricultural; unmanned aerial vehicle; stress resistance; virtual reality; Unreal Engine; gaze tracking; mission planning.

Citation: Taskina LA, Abakumov LA, Kazarkin TD. A method for analyzing biometric data to assess the cognitive load and stress resistance of an agricultural unmanned aerial vehicle operator during mission planning. *Computer Optics* 2025; 49(6): 1202-1210. DOI: 10.18287/COJ DOI: 10.18287/COJ1849.

Introduction

Mission planning for agricultural unmanned aerial vehicles (UAV) is a cognitively demanding perceptual task in which the operator must account for geofences, above ground level (AGL) and mean sea level (MSL), energy constraints, weather, safety and connectivity requirements [1–4]. Errors at this stage lead to costs and risks; traditional checklist-style assessments do not reflect the operator's internal state, including overload, fatigue, or stress [5–7]. Modern virtual reality (VR) systems with integrated eye tracking provide objective signals – where and how the operator looks – enabling pose estimation and fixation detection, blink and eyelid openness, and pupil diameter [8]. These variables correlate with workload and stress. However, industrial application programming interfaces (APIs), such as OpenXR and PICOXR, primarily expose aggregated indicators (gaze rays, openness, etc.), while raw infrared (IR) eye images are not available for privacy reasons. This can hinder the use of the rich computer vision toolset for pupil and eyelid analysis [9, 10].

Eye-derived features like blink frequency, duration, pupil dynamics have been used previously for fatigue analysis [11–13]. In VR and human-computer interaction, gaze-to-object mapping, heatmaps, and scanpath analysis are well studied [14]. More recently, ML and DL approaches (LSTM, TCN, Transformer) exploited temporal models and image-like attention representations such as recurrence heatmaps to classify state patterns [15–17]. Eye tracking plays a special role in analyzing interaction with interfaces [18–20], where the tracker enables geometrically precise binding of gaze to user interface (UI) elements—for example, during UAV mission planning [21]. In practice, missing visual biological cues often need to be substituted with informative behavioral surrogates (area of interest (AOI) transition entropy, reaction latencies, fixation patterns) when raw eye images are unavailable [18]. It is also essential to ensure balanced data and sufficient measurements of rare stress episodes across participants to obtain reliable, robust models and avoid overfitting [22].

In this work we propose a vision-oriented pipeline in which precise gaze-to-UI geometry yields metrically correct recurrence-heatmap images and rich AOI-graph analytics. We show how aggregated eye signals can be converted into robust features of workload and stress for agricultural UAV mission planning without access to raw frames. We present results for training a hybrid model comprising a temporal TCN [23] and an image-based convolutional neural network (ResNet-18) [24], validated in a subject-out regime. A key aspect is synthetic data augmentation to improve robustness given the small number of stress episodes.

1. Approaches and methods

Hardware. We use a PICO 4 Enterprise VR headset with integrated eye tracking to present a virtually modeled mission-planning UI. Eye data were sampled at 90 Hz and synchronized with the VR frame rate. Access to biometrics was provided via the PICOXR plugin exposing aggregated signals: the gaze ray (origin and direction) both combined and

per-eye, eyelid openness metrics and blink flags, and pupil diameter. Raw IR frames are unavailable per the software development kit’s (SDK) security model, hence the method focuses on processing aggregates (gaze, blink, pupil) and on constructing image-like attention representations from geometrically accurate gaze hits on UI elements [9, 10].

Software. The system is implemented in Unreal Engine 5. The mission-planning scene is a world-anchored set of widgets:

1. MapCanvas – main workspace with a base map and drawn route polylines;
2. Waypoint – route waypoint panel;
3. Validation – warnings such as no-fly intersections, AGL-min violations, low battery;
4. Constraints – mission rule constraints and settings;
5. Toolbar – primary mission commands and tools.

As a reference mission-planning UI we adopted QGroundControl [21], widely standardized and used in real agricultural UAV missions. The virtual UI is arranged to avoid head-locked window artifacts, yielding sTab. Areas of Interest (AOIs) for attention analytics. Widgets are placed within the working volume 0.9–1.1 m from the user, with a fixed orientation toward the head-mounted display (HMD) and a “dead zone” for comfort. A ray pointer from the controller is used for direct interaction. Each session starts with the HMD’s standard eye tracking calibration, followed by the planning scenario: the operator places or edits waypoints, sets AGL/MSL, draws no-fly polygons, runs optimization and validation; when the stress-test scenario is activated, the operator performs replanning. Fig. 1 shows the agricultural UAV mission-planning UI.



Fig. 1. Structure of widgets for agricultural UAV mission planning

Gaze-to-UI geometry. The key component is geometric intersection of the gaze ray with UI elements. For each frame, PICOXR yields a pair (o, d) – ray origin and direction. Widget planes in UE5 are described by a point p_0 and normal n . The intersection point is $h = o + td$, where

$$t = n \times (p_0 - o) / n \times d \tag{1}$$

The point h is then affinely projected to UV coordinates:

$$(u, v) \in [0, 1]_2 \tag{2}$$

to look up the AOI mask. This yields metrically correct time series of AOI hits and enables constructing “attention images,” heatmaps, and recurrence plots without access to raw eye frames. To mitigate microsaccades and head jitter we smooth gaze directions (50–100 ms sliding window) and drop samples with invalid tracking (PICOXR validity flags).

Features. Two feature families are derived. Behavioral-geometric features are computed from AOI-hit sequences and ray kinematics: dwell fractions per AOI, AOI transition entropy, mean and distribution of fixation durations (thresholded by ray speed), number of regressions, reaction latency to salient events (e.g., time from new no-fly appearance to first corrective action). Eye aggregates include blink rate and duration, openness profile, and pupil size: change amplitude, area under the curve, and recovery time after “Validate” and “Optimize” events [11–13].

In parallel we build image-like attention representations. For each window, points (u_i, v_i) are converted into a heatmap H via kernel density and into a recurrence plot R reflecting revisits. These images are further masked by binary AOI maps, facilitating subsequent interpretation of attention by UI entity (map, constraints panel, waypoint list, etc.). Thus, even from aggregated tracking, we obtain a full-fledged vision input for convolutional models.

Model. The training architecture has two branches with late fusion. The temporal branch processes time series (u_i, v_i) , their velocities, blink and openness, and UI event markers using a TCN [23]. TCN configuration: window 30 s, hop 15 s, sampling 30 Hz (downsampled from 90 Hz); five residual blocks with dilations {1, 2, 4, 8, 16}, kernel size 5, channels [64, 64, 128, 128, 256], ReLU activations. We use WeightNorm, dropout = 0.1, AdamW (lr = 3e-4, weight

decay = 0.05), and early stopping under subject-out validation [22]. The image branch ingests H and R and extracts embeddings via a lightweight ResNet-18 [24], with 128×128 inputs and the first convolution adapted to two channels (initialized by the mean of RGB weights). At fusion, vector representations are concatenated and fed into heads for (a) multi-class classification of cognitive load (L1–L3) and (b) binary stress resistance estimation (small vs. large quality drop in stress scenarios). To curb overfitting with limited, imbalanced data, we apply contrastive pretraining of the temporal branch on synthetic gaze trajectories (generator controls saccade amplitude and matches the power spectrum of gaze velocity), and class balancing with tabular synthetic data via CTGAN [25] for windows with rare events. Critically, all synthetic data are used only for training; validation follows a strict subject-out protocol on real data [22].

Mission-planning performance metrics. Key performance indicators are computed directly in the VR scene. We log planning time T , time-to-replan TTR , number of conflicts, and number of edits. Route quality is evaluated by

$$J = \alpha L + \beta \int risk(s) ds + \gamma [penalty]_{AGL}, \quad (3)$$

where L is the path length; the risk integral accounts for no-fly intersections and penalties for undesirable altitude profiles. Coefficients α, β, γ are chosen per scenario assumptions. These metrics are analyzed jointly with gaze-AOI features and model predictions to (1) detect rising cognitive load in stress scenes and (2) compute an integrated stress resistance index based on normalized losses in J and T , recovery speed after an event, and viewing-strategy orderliness (AOI transition entropy).

Logging and synchronization. Reliable logging and temporal synchronization are essential. All eye and UI events are recorded with timestamps shared by the renderer and sensors. Logs include: gaze ray, validity and AOI hits, openness and blink, pupil size, UI events (add/move/delete waypoint, optimization, validation, no-fly input/appearance, timer start/expiration), and route parameters. This audit trail enables reproducible feature computation, model training, and ex-post verification without device access.

In sum, the combined hardware-software setup and geometrically rigorous gaze mapping circumvent the lack of raw eye frames by relying on sTab. aggregated channels and attention visualizations as full-value computer vision inputs.

2. Experimental setup and data collection

Scenes and Participants. The study comprised four mission-planning scenes in the UI that differed in complexity and stressors:

- 1.L1 (baseline): route layout with a given start point, one no-fly zone, a minimum AGL, a battery constraint, and no timer; duration 6 – 8 min;
- 2.L2 (extended constraints): multiple no-fly zones, altitude corridors, a weather module as an additional risk layer; duration 8 – 10 min;
- 3.L2–L3 (replanning): after a successful Validate, a new no-fly zone is injected into the scene and replanning is required; the injection occurs at $t=3$ after the “Validate” button press.
- 4.L3 (timer): L2 complexity with a countdown $T=15$ s and audio cues; when time expires, the mission is deemed incomplete.

In each scene, the operator followed the sequence: place or correct waypoints → set AGL/MSL → Optimize → Validate → Export. Twelve volunteers with no acute vestibular complaints participated; all provided informed consent and agreed to the experimental terms. Scene order was counterbalanced. Prior to the main session, eye tracking calibration and a 3–5 min training scene were performed. Gaze calibration followed the HMD’s standard procedure until validity $\geq 95\%$ on test markers was achieved. During the session, each gaze sample carried a validity flag; samples marked Invalid and gaps > 200 ms were excluded. To stabilize the track, exponential smoothing of the gaze direction with a 50–100 ms window was applied.

Gaze-to-UI and AOI annotation. As described in Section 1, at each frame we computed the intersection point h of the gaze ray with the plane of the corresponding widget, then affinely projected it into UV coordinates and classified it using AOI masks. This produced precise time series of hits on functional UI zones. To bind gaze to action, we introduced an event anchor t_{evt} (e.g., a button press or waypoint grab). A hit was considered correct if the last valid gaze sample within $[t_{evt} - 100 \text{ ms}, t_{evt} + 50 \text{ ms}]$ fell inside the target AOI. This rule was used both to compute AOI-hit accuracy on interactions and to monitor mapping quality.

Data Quality and Missingness. For each scene we assessed the fraction of valid gaze samples near events and the stability of fixations (standard deviation of angular jitter). Validity levels were: L1 92–96%, L2 90–95%, L2–L3 87–93%, L3 82–90%; in the stress scene, reduced validity was accompanied by increased jitter. Across all participants and scenes, 1198 UI events were recorded, of which 1023 resulted in correct hits on target AOIs (aggregated AOI-hit accuracy $\approx 85.3\%$). A detailed breakdown by interaction type is provided in Tab. 1. These values serve both as a report on Gaze-to-UI mapping quality and as empirical guidance for subsequent processing.

Dataset Construction. From the continuous stream, we derived tabular behavioral-geometric features: time fractions per AOI, AOI transition entropy, fixation and regression parameters, reaction latencies, and the dynamics of openness, blink or pupil. We also constructed image-like attention representations: heatmaps and recurrence plots of AOI visits in sliding windows of 30–60 s (50% overlap). For learning tasks, data were split subject-out; participants did not overlap between train and test. Feature normalization was performed on the training portion of each fold. Thus privacy and

reproducibility were ensured: public artifacts (UI illustrations, AOI masks, heatmaps, and schematics) are reproducible from logs without device access and contain no personally identifiable information.

Synthetic Data Supplementation. Real stress episodes during mission planning are rare and heterogeneous in duration. To avoid overfitting and to balance classes (L1–L3 and stress vs. non–stress), we augmented the real dataset with two types of synthetic samples used only during training: (a) temporal sequences of gaze and events, and (b) attention images derived from those sequences.

Parametric Generator of Oculomotor Activity. The base gaze stream (u_t, v_t) was constructed as alternating fixations and saccades (Fig. 2):

1. Fixations: durations $T_{fix} \sim \text{LogNormal}(\mu, \sigma)$, drift modeled as angular white noise with scene–dependent variance;
2. Saccades: amplitudes from an exponential distribution (e.g., log–normal or Pareto); directions drawn from a mixture of a uniform and a goal–directed component (toward target AOIs). Duration and peak velocity obey the main sequence relationship with amplitude;
3. Blinks and validity: a Poisson process of abrupt validity dropouts, with higher probability in stress scenes;
4. Event–conditioned behavior: upon appearance of a new no–fly zone, the probability of gaze transitions to the Map AOI and Constraints panel increases; around Validate, transitions to the Toolbar and Validation panel and back to the Map become more likely.

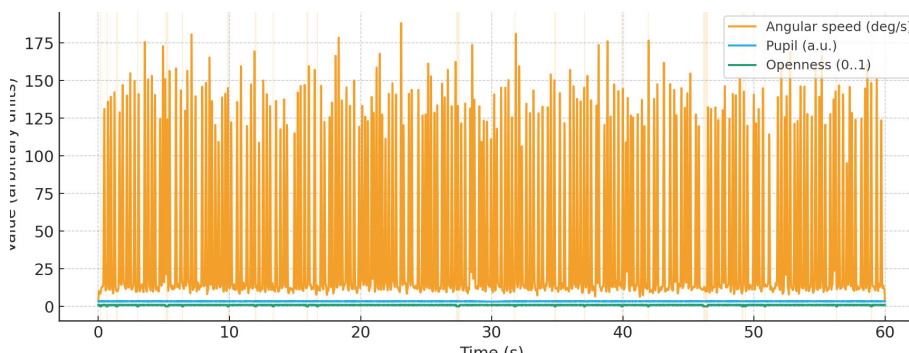


Fig. 2. Structure of the generated synthetic oculomotor time series: angular gaze velocity, pupil, and openness traces

AOI Transition Model. The AOI sequence was encoded as a discrete process with a transition matrix $p^{(s)}$ for each scene s (L1...L3). $P^{(s)}$ was estimated from real data and modified per class (e.g., in L3, increased transitions Map → Constraints, Map → Validation, and a higher share of short fixations). This enforces realistic viewing strategies.

Temporal Data augmentation. On the synthetic series we applied mild deformations such as time–warp ($\pm 5–10\%$), UV jitter, and the insertion of rare mini–dropouts in validity – bounded so as not to break gaze–to–UI geometry. From the synthetic series we then produced kernel–density heatmaps H of (u_t, v_t) with AOI masking, and binary recurrence plots of AOI revisits over a 30–60 s window, mirroring the real–data pipeline (an example synthetic map is shown in Fig. 3). To balance classes, we generated stress windows with more fragmented fixations and elevated AOI transition entropy.

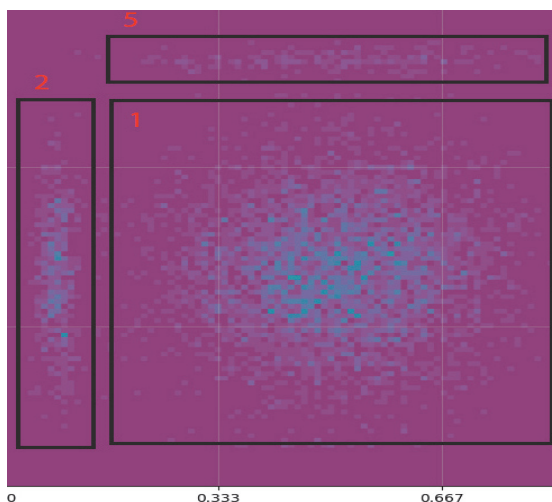


Fig. 3. Example synthetic UV heatmap with AOI outlines: 1–MapCanvas, 2–Waypoints, 3–Validation, 4–Constraints, 5–Toolbar

Balancing and Proportions of Synthetic vs. Real Data. To equalize rare metric combinations e.g., high reaction latency + low AOI–hit + high risk J we synthesized tabular feature windows as fixation statistics and AOI entropies using CTGAN, with constraints enforcing physical plausibility. The CTGAN generator used multilayer perceptrons (4 layers \times 256 units, LeakyReLU) and was trained with WGAN–GP and PacGAN (pack = 2).

Mini-batches mixed synthetic and real examples at 1:1 for rare classes and 1:3 for frequent classes. Synthetic data were generated only from statistics of the current fold’s training subjects. We performed duplicate checks via nearest neighbors in embedding space to filter near-copies of real windows. Training on synthetic data proceeded in two stages: first on real + mild synthetic (without extreme jitter or dropouts), then progressively adding hard variants with stress noise and dense missingness.

In this way, the generation-and-verification loop safely expands the training set while preserving the realism of oculomotor dynamics and UI interaction patterns, and it controls statistical proximity to real data. As a result, the target models more robustly detect increases in cognitive load and decreases in stress resistance in rare stress scenes without degradation on real subject-out validation. In sum, the described hardware-software configuration, rigorous geometric gaze-to-UI alignment, synthetic data supplementation, and disciplined logging yield a synchronized dataset suiTab. for computing cognitive-load and stress resistance metrics in agricultural UAV mission-planning scenarios.

3. Results

Gaze-to-UI accuracy on interactions. We measured gaze hits within target AOIs for all interaction events (Tab. 1). The share of interactions counted toward AOI-hit accuracy-button press and release (Validate, Optimize), map click, waypoint drag start and end, list-row selection-was computed as:

$$Accuracy = (number\ of\ correct\ hits) / (total\ interactions) \times 100\ \% \tag{4}$$

A gaze was counted as a hit if, within $[t_{event} - \Delta t_{pre}, t_{event} + \Delta t_{post}]$, the last valid gaze sample fell inside the target AOI.

Tab. 1. Reliability of gaze-to-UI mapping: AOI-hit accuracy by scene and interaction type

Scene	Interaction type	N events	Correct hits	AOI-hit Accuracy, %	95 % CI Wilson	Valid gaze rate, %
L1	Click Map (AddWP/Move)	226	202	89.4	84.7–92.8	94
	Button Validate	36	32	88.9	74.7–95.6	96
	Waypoint list select	60	52	86.7	75.8–93.1	93
L2	Click Map (Add/Move)	264	228	86.4	82.1–90.3	92
	Button Validate	48	42	87.5	75.3–94.1	94
	Constraints toggle	48	40	83.3	70.4–91.3	90
L2-L3	Drag WP (replan)	144	121	84.0	77.2–89.1	90
	Button Validate	36	31	86.1	71.3–93.9	92
L3	Click and Drag under timer	288	236	81.9	77.1–86.0	86
	Button Validate	48	38	79.2	65.7–88.3	88
All	All types	1198	1023	85.3	83.3–87.3	90

In total, 1198 UI events were recorded, of which 1023 were correct AOI hits (AOI-hit accuracy = 85.3 %, see Tab. 1). Accuracy decreased monotonically with increasing scene difficulty and stress. Tab. 1 also shows a systematic decline in the fraction of valid samples in the event-synchronization window: L1 = 94.1 %, L2 = 92.0 %, L2-L3 = 90.4 %, L3 = 86.3 %. By interaction type, buttons (e.g., Validate) expectedly yield higher hit accuracy than map actions (Click and Drag). However, in the L3 stress scene even buttons show a drop (validate in L3 = 79–88 %).

AOI transition matrices. Fig. 4–5 illustrate viewing strategies across AOIs as stress increases.

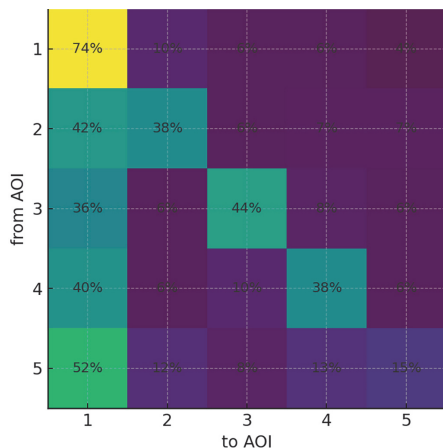


Fig. 4. AOI transition matrix in scene L1 (row-normalized)

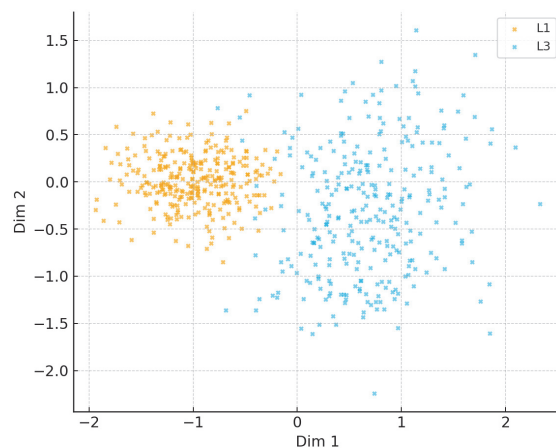


Fig. 5. AOI transition matrix in scene L3 (row-normalized)

In Fig. 4, prolonged dwell on MapCanvas (1) dominates, with moderate excursions to Waypoints (2) and Toolbar (5). In the more complex scene (Fig. 5), shuttle-like transitions intensify between Map (1) → Constraints (4) and Map (1) → Validation (3), while long, monotonic fixations on the map decrease. With a timer (L3), this effect is strongest and aligns

with more frequent constraint checks and attempts to “beat” the deadline. The growth of these transitions is accompanied by higher transition entropy, as shown in Fig. 6.

Fixations and search orderliness. Fixation–duration distributions shorten with increasing load: the median shift from L1 to L3 reflects a rougher strategy with frequent saccades (Fig. 7). The AOI transition–entropy curve rises from L1 to L3 (Fig. 6), consistent with fragmented attention under constraints and timer pressure.

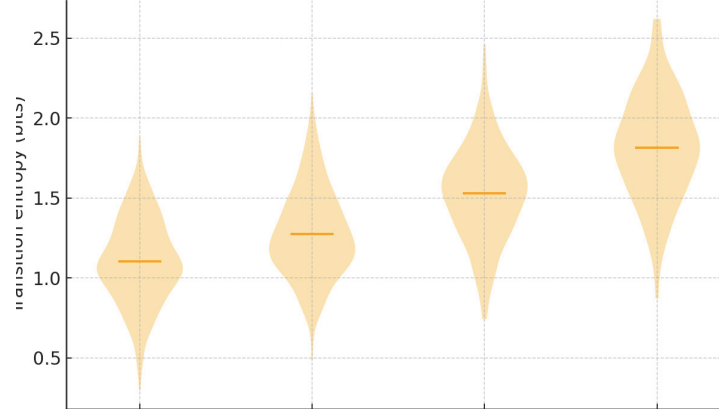


Fig. 6. AOI transition entropy by scene L1–L3

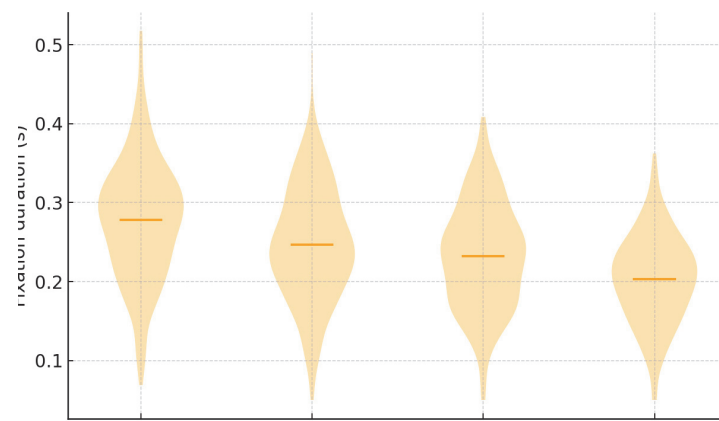


Fig. 7. Fixation duration distributions by scene L1–L3

UV–domain heatmaps for real windows and synthetic controls show concentration in AOIs 3–4 in complex scenes and an attention shift from Waypoints (2) toward Constraints and Validation during replanning. The recurrence plot confirms more frequent returns to the same regions in L3, consistent with the transition matrices.

Event–locked analyses around Validate and the appearance of a new no–fly zone show a larger and longer pupil deflection and a deeper dip in L3 compared to L1. Qualitatively, this indicates stronger physiological responses to control/error checks and reduced fixation stability under load. Overall, event amplitude (peak–to–baseline) and recovery time are larger in L3 (slower return to baseline), matching the rise in AOI transition entropy.

Classification of load and stress resistance. Behavioral metrics collected directly in VR show expected trends:

1. Planning time T increases from L1 to L2 and L3;
2. Time–to–replan (TTR) after no–fly injection in L2–L3 shortens for some participants (reflexive dragging), while edits and UI conflicts increase;
3. The route–quality functional J degrades in L3, indicating a tight speed–quality trade–off under a timer.

Window–level correlation analysis shows that AOI entropy and latency of the first corrective action are positively associated with degradation in J and with lower AOI–hit; the more fragmented the viewing and the slower the “first step” after an event, the worse the route and interaction accuracy. These associations strengthen in L2–L3 and L3.

The trained hybrid model TCN on sequences plus ResNet–18 on attention images consistently outperforms single–branch baselines for both load level (multiclass L1, L2, L2–L3 and L3) and stress resistance (binary degradation under stress). Under subject–out evaluation, receiver operating characteristic (ROC) and precision–recall (PR) curves show confident class separability (Figs. 8–9).

Ablations quantify each branch’s contribution: the TCN branch captures temporal structure (fixations, saccades, latencies), while the convolutional neural network branch captures spatial attention patterns. Adding synthetic windows to training improves recall on rare classes (L2–L3 and L3) without harming real–only subject–out test performance. Fig. 10 shows 2D projections of embeddings: L1 windows form compact clusters, whereas L3 splits into multiple subclusters with higher variability, consistent with rising transition entropy and reduced fixation stability.

Conclusion

We presented a vision-oriented method for assessing cognitive load and stress resistance of agricultural UAV operators during mission planning in VR, based on geometrically rigorous gaze-to-UI mapping, subsequent AOI graph analytics, and image-like attention representations. Experiments across four scenes (L1-L3) show that:

1. Mapping is reliable, with AOI-hit 85 % over 1198 events (Tab. 1). The valid-gaze rate declines from L1 to L3 yet remains sufficient for analysis;
2. Behavioral and ocular features respond systematically to rising complexity and stress: T and TTR increase, AOI transition entropy rises, event-locked pupil responses appear, and brief openness drops occur. L3 transition matrices indicate a loosening of viewing strategy (Figs. 4–5);
3. The subject-out-trained hybrid model outperforms single branches; fold-wise ROC and PR curves show confident separability of load classes (Figs. 8–9). Embedding structure illustrates compact L1 clusters and L3 splitting (Fig. 10). Synthetic augmentation of gaze trajectories and tabular windows improves robustness to rare stress episodes without degrading validation;
4. The method does not require raw IR eye-camera frames, so all computations rely on SDK aggregates and attention images important for privacy and portability.

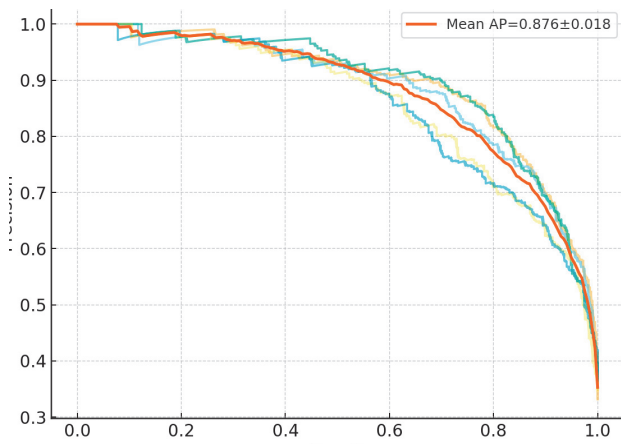


Fig. 8. Precision-recall curves across folds

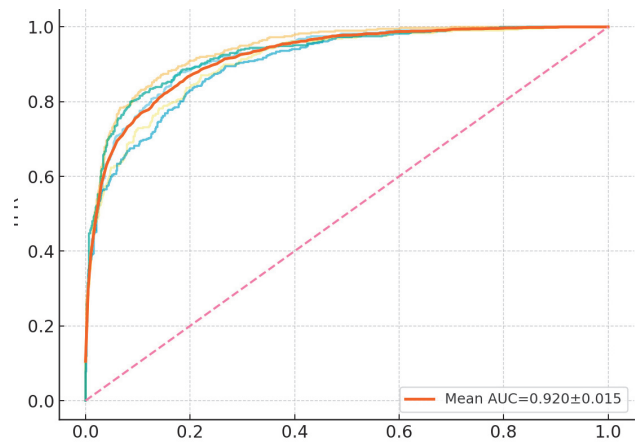


Fig. 9. ROC curves across subject-out folds

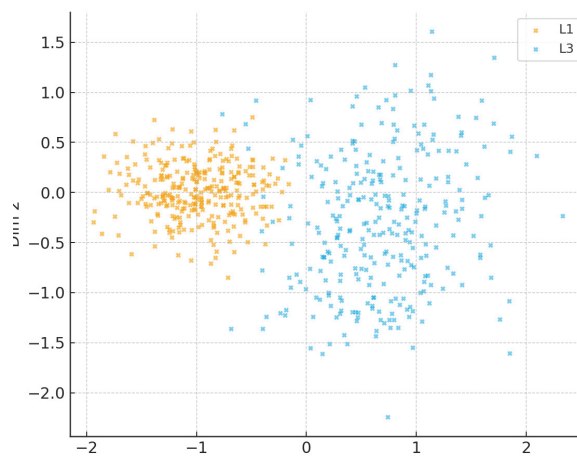


Fig. 10. 2D projection of fused embeddings

Limitations and future work. Sample size $N=12$ and a limited spectrum of UI tasks. In some stress scenes the valid-rate drops, necessitating robust filtering rules and extended calibration (Tab. 1). Future work will address:

1. expanding the task set and transferring to real ground control station, such as QGroundControl and Mission Planner;
2. integrating load assessment with adaptive UI hints;
3. richer synthetic data (GAN processes for saccades, agent-based attention simulators);
4. testing generalization across other HMDs and SDKs.

Acknowledgments

This work was supported by the Ministry of science and higher education of the Russian Federation, grant No 075–15–2025–610.

References

- [1] Meng W, Zhang X, Zhou L, Guo H, Hu X. Advances in UAV path planning: A comprehensive review of Methods, Challenges, and Future Directions. *Drones* 2025; 9(5): 376. DOI: 10.3390/drones9050376.
- [2] Kim J, Atkins E. Airspace Geofencing and Flight Planning for Low-Altitude, Urban, Small Unmanned Aircraft Systems. *Applied Sciences* 2022; 12(2): 576. DOI: 10.3390/app12020576.
- [3] ElSayed M, Mohamed M. Robust digital-twin airspace discretization and trajectory optimization for autonomous unmanned aerial vehicles. *Scientific Reports* 2024; 14: 12506. DOI: 10.1038/s41598-024-62421-4.
- [4] Behjati M, Nordin R, Zulkifley MA, Abdullah NF. 3D Global Path Planning Optimization for Cellular-Connected UAVs under Link Reliability Constraint. *Sensors* 2022; 22(22):8957. DOI: 10.3390/s22228957.
- [5] Das Chakladar D, Roy PP. Cognitive workload estimation using physiological measures: a review. *Cogn Neurodyn* 2024; 18(4):1445-1465. DOI: 10.1007/s11571-023-10051-3.
- [6] Radhakrishnan V, Louw T, Gonçalves RC, Torrao G, Lenné M, Merat N. Using pupillometry and gaze-based metrics for understanding drivers' mental workload during automated driving. *Transportation Research Part F: Traffic Psychology and Behaviour* 2023; 94: 254-267. DOI: 10.1016/j.trf.2023.02.015.
- [7] Alshanskaia EI, Portnova GV, Liaukovich K, Martynova OV. Pupillometry and autonomic nervous system responses to cognitive load and false feedback: An unsupervised machine learning approach. *Frontiers in Neuroscience* 2024; 18:1445697. DOI: 10.3389/fnins.2024.1445697.
- [8] Ugwitz P, Kvarda O, Juříková Z, Šašinka Č, Tamm S. Eye-tracking in interactive virtual environments: Implementation and evaluation. *Applied Sciences* 2022; 12(3):1027. DOI: 10.3390/app12031027.
- [9] Warin C, Pak V, Reinhardt D. Privacy perceptions across the XR spectrum: An extended reality cross-platform comparative analysis of a virtual house tour. *Proceedings on Privacy Enhancing Technologies* 2025; 1: 150. DOI: 10.56553/popets-2025-0009.
- [10] Garrido GM, Nair V, Song D. SoK: Data privacy in virtual reality. *Proceedings on Privacy Enhancing Technologies* 2024; 21: 40. DOI: 10.56553/popets-2024-0003.
- [11] Kuwahara A, Nishikawa K, Hirakawa R, Kawano H, Nakatoh Y. Eye fatigue estimation using blink detection based on Eye Aspect Ratio Mapping (EARM). *Cognitive Robotics* 2022; 2: 50-59. DOI: 10.1016/j.cogr.2022.01.003.
- [12] Kaur K, Gurnani B, Nayak S, Deori N, Kaur S, Jethani J, Singh D, Agarkar S, Hussaindeen JR, Sukhija J, Mishra D. Digital eye strain- A comprehensive review. *Ophthalmol Ther.* 2022; 11(5):1655-1680. DOI: 10.1007/s40123-022-00540-9.
- [13] Huyghe T, Calleja-González J, Bird SP, Alcaraz PE. Pupillometry as a new window to player fatigue? A glimpse inside the eyes of a Euro Cup Women's Basketball team. *Biol Sport.* 2024; 41(1): 3-15. DOI: 10.5114/biolSport.2024.125590.
- [14] Bisogni C, Nappi M, Tortora G, Del Bimbo A. Gaze analysis: A survey on its applications. *Image and Vision Computing* 2024; 144(13): 104961. DOI: 10.1016/j.imavis.2024.104961.
- [15] Szczepaniak D, Harvey M, Deligianni F. Predictive Modelling of Cognitive Workload in VR: An Eye-Tracking Approach. *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications* 2024; 46: 1-3. DOI: 10.1145/3649902.3655642.
- [16] Moulder R, Booth B, Abitino A, D'Mello S. Recurrence Quantification Analysis of Eye Gaze Dynamics During Team Collaboration. *ACM LAK* 2023. DOI: 10.1145/3576050.3576113.
- [17] Atweh J, Riggs S. Quantifying collaborative strategies and identifying performance breakdowns of UAV C2 teams using multidimensional cross-recurrence quantification analysis. *International Journal of Human-Computer Studies* 2025; 204: 103593. DOI: 10.1016/j.ijhcs.2025.103593.
- [18] Wu S, Chen H, Hou L, Zhang GK, Li CQ. Using Eye-Tracking to Measure Worker Situation Awareness in Augmented Reality. *Automation in Construction* 2024; 165: 105582. DOI: 10.1016/j.autcon.2024.105582.
- [19] Devlin SP, Brown NL, Drollinger S, Sibley C, Alami J, Riggs SL. Scan-based eye tracking measures are predictive of workload transition performance. *Applied Ergonomics* 2022; 105: 103829. DOI: 10.1016/j.apergo.2022.103829.
- [20] Zhu M, Wu Q, Bai Z, Song Y, Gao Q. EEG-eye movement based subject dependence, cross-subject, and cross-session emotion recognition with multidimensional homogeneous encoding space alignment. *Expert Systems with Applications* 2024; 251: 124001. DOI: 10.1016/j.eswa.2024.124001.
- [21] Ramírez-Atencia C, Camacho D. Extending QGroundControl for automated mission planning of UAVs. *Sensors* 2018; 18(7): 2339. DOI: 10.3390/s18072339.
- [22] He M, Alkurdi A, Clore JL, Sowers RB, Hsiao-Weckler ET, Hernandez ME. Scoping Review of ML Approaches in Anxiety Detection from In-Lab to In-the-Wild. *Applied Sciences.* 2025; 15(18): 10099. DOI: 10.3390/app151810099.
- [23] Bednarski BP, Singh AD, Zhang W, Jones WM, Naeim A, Ramezani R. Temporal convolutional networks and data rebalancing for clinical length of stay and mortality prediction. *Scientific Reports* 2022; 12: 21247. DOI: 10.1038/s41598-022-25472-z.
- [24] Shafiq M, Gu Z. Deep Residual Learning for Image Recognition: A Survey. *Applied Sciences* 2022; 12(18): 8972. DOI: 10.3390/app12188972.
- [25] Xu L, Skoularidou M, Cuesta-Infante A, Veeramachaneni K. Modeling tabular data using conditional GAN. *NeurIPS Workshop* 2019; 659: 7335. DOI: 10.48550/arXiv.1907.00503.

About authors

Taskina Larisa Anatolyevna, (b. 1991) graduated from Samara National Research University in 2015 with a Master's degree, majoring in Applied mathematics and physics. Now she is a Junior Research Fellow at the World-Class Scientific Center "Intelligent Unmanned Aircraft Systems". Research interests: synthetic data generation, virtual reality, computer image processing, data science and deep learning. E-mail: lara.zherdeva.taskina@gmail.com

Abakumov Leonid Alexandrovich, (b. 2001) graduated from Samara National Research University in 2023 with a Bachelor's degree, majoring in Applied Mathematics and Computer Science. Now he is a Master's student at Samara University and a laboratory assistant at Research Center "Intelligent mobility of multifunctional unmanned aircraft systems". Research interests: virtual reality, pattern recognition, synthetic data, computer image processing, data science and deep learning. E-mail: leonid6845@mail.ru

Kazarkin Timofey Dmitrievich, (b. 2001) graduated from Samara National Research University in 2023 with a Bachelor's degree, majoring in Applied Mathematics and Computer Science. Now he is a Master's student at Samara University and a laboratory assistant at Research Center "Intelligent mobility of multifunctional unmanned aircraft systems". Research interests: virtual reality, pattern recognition, synthetic data, computer image processing, data science and deep learning. E-mail: td.kazarkin@gmail.com

Received September 29, 2025. The final version – November 16, 2025 г.
