Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 134–144
https://doi.org/10.54941/ahfe1008248

Physiological Indicators of Trust in AI-Assisted Air Traffic Control: A Multimodal Eye-Tracking and EEG Study

Wenqian Qi[1], Xing Yao[2], Bufan Liu[2], Xinyu Liu[3], Hu Li[2], Tianjiao Zhang[4], and Xiaoqing Yu[2]

[1]School of Art and Design, Wuhan University of Technology, Wuhan, 430070, Hubei, China
[2]School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798
[3]School of Information, Yunnan University, Kunming, 650500, Yunnan, China
[4]College of Information Technology, Shanghai Ocean University, Shanghai 201306, China

ABSTRACT

The increasing use of artificial intelligence (AI) in air traffic management requires operators to maintain appropriate trust in automated decision-support systems. Excessive trust may lead to insufficient monitoring, whereas insufficient trust may result in system disuse and increased workload. This study investigates whether eye-tracking and electroencephalography (EEG) signals can serve as objective indicators of trust during a simulated AI-assisted air traffic control task. Nineteen participants completed a simulated conflict-detection and resolution task with an AI advisor operating at low, high, and perfect reliability levels. After data-quality screening, 2,536 valid trials were retained. Eye-tracking and EEG signals, behavioural responses, and subjective trust ratings were recorded, and physiological features were subsequently extracted for classification. Trust classification was performed using Logistic Regression, Support Vector Machine, Random Forest, XGBoost, and a one-dimensional Convolutional Neural Network. Random Forest achieved the best overall performance, with eye-tracking features providing the highest classification accuracy ($F1 = 0.759 \pm 0.059$). Low-trust trials were associated with longer fixation durations, greater pupil dilation, and elevated or less stable beta-band EEG activity. These findings demonstrate the feasibility of using non-invasive physiological sensing for trust monitoring in safety-critical human–AI collaboration. The results further suggest that robust and interpretable eye-tracking features may provide a practical foundation for real-time trust-aware adaptive decision-support systems.

Keywords: Trust in automation, Eye tracking, Electroencephalography, Air traffic control, Human–AI collaboration, Neuroergonomics

INTRODUCTION

The rapid development of intelligent transportation systems has accelerated the adoption of intelligent decision-support technologies across transportation domains (Sarker et al., 2020; Yao et al., 2022). Within aviation, AI has been increasingly applied to both operational decision support and post-accident

Received August 11, 2026; Revised September 26, 2026; Accepted October 2, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

human-factors analysis (Liu et al., 2025). In air traffic management, automation and AI support operators in traffic monitoring, conflict detection, and time-critical decision-making (Lyu and Li, 2026; Xia et al., 2026; Yu et al., 2025; Yu et al., 2023). Automated aids such as Short-Term Conflict Alert and Medium-Term Conflict Detection can identify potential losses of separation and provide warnings or recommended actions (Bestugin et al., 2020). As these systems become more capable and autonomous, appropriate trust becomes essential for effective and safe human–automation interaction.

Trust in automation refers to operators' willingness to rely on automated recommendations under uncertainty. Appropriate trust calibration reduces workload and improves decision-making, whereas over-trust may cause complacency and under-trust may increase unnecessary monitoring (Hoff and Bashir, 2015). Trust is commonly assessed using questionnaires and rating scales. Although subjective measures are straightforward to administer, they interrupt task performance and provide only discrete rather than continuous observations of trust (Kohn et al., 2021). Physiological sensing offers a complementary approach because it can capture objective and time-varying responses without repeatedly interrupting the operator. Eye tracking can quantify visual monitoring behaviour (Li et al., 2025), attention allocation (Xia et al., 2025b), fixation patterns, and pupil responses (Lyu and Li, 2025; Yao et al., 2026a), while electroencephalography can provide information about neurophysiological activity associated with cognitive processing and system evaluation (Hergeth et al., 2016; Oh et al., 2020). Previous studies have demonstrated the feasibility of predicting trust using EEG, eye tracking, galvanic skin response, and multimodal physiological signals (Akash et al., 2018; Yao et al., 2026b). Nevertheless, much of this work has focused on automated driving, unmanned systems, investment games, or abstract monitoring tasks. Comparatively fewer studies have examined trust using multimodal physiological measurements in realistic air traffic management scenarios.

This study addresses this gap by investigating trust-related eye-movement and EEG patterns during a simulated ATC conflict-monitoring task supported by an AI advisor of varying reliability. The study has two objectives: (1) to evaluate whether eye-tracking and EEG features can distinguish low and high subjective trust; (2) to compare eye-tracking, EEG, and multimodal input for trust classification.

RELATED WORKS

Trust calibration refers to the degree to which operator trust corresponds to the actual capability and reliability of an automated system (Chiou and Lee, 2023). Calibrated trust supports appropriate reliance, while miscalibrated trust may lead to misuse or disuse. Although trust is influenced by individual, cultural, organisational, and contextual factors, automation performance remains a central determinant (Hoff and Bashir, 2015). Failures can affect trust asymmetrically. Negative events may reduce trust more quickly than successful performance restores it. The relative effects of misses and false alarms are task-dependent. In automated-driving contexts, misses may be regarded as more harmful because an undetected hazard could have severe

consequences (Clark et al., 2010). In monitoring tasks, false alarms may be especially disruptive because operators must redirect attention and investigate an event that does not require action. These findings highlight the need to evaluate automation failures within the specific operational context in which the automation is used.

Self-report scales remain the most common measures of trust in human–automation interaction (Kohn et al., 2021). However, subjective reports are retrospective, may vary across individuals, and cannot continuously represent changing trust states. EEG provides a non-invasive measure of neural activity and has been used to identify cognitive responses to automated systems (Li et al., 2023; Yu et al., 2026; Yu et al., 2025). EEG has been widely used to investigate trust-related neural activity, although reported associations between frequency bands and trust vary across experimental contexts (Akash et al., 2018; Oh et al., 2020; Zhang et al., 2024). This variation suggests that EEG markers of trust are likely context-dependent and should be interpreted in relation to task demands. Eye tracking offers an alternative and comparatively practical means of capturing various cognitive states such as fatigue (Yu et al., 2023), stress (Xia et al., 2025a), task load (Liu et al., 2025; Liu et al., 2024) and trust-related monitoring behaviour (Yao et al., 2026b). Higher trust has been associated with less frequent monitoring and shorter fixation durations, whereas lower trust may require operators to spend more time checking or verifying automated recommendations (Körber et al., 2018; Walker et al., 2019). Longer fixations and more frequent monitoring have also been linked to successful automation-failure detection (Bruder and Hasse, 2020). Pupil diameter may further reflect changes in cognitive effort, risk processing, and task involvement. Multimodal sensing may capture complementary aspects of trust. Previous studies have reported that combining physiological signals can improve predictive performance compared with a single modality (Jiang et al., 2026). However, multimodal fusion also increases sensor burden, preprocessing complexity, and the possibility of adding noisy or redundant information. Therefore, it is important to empirically compare multimodal and unimodal approaches rather than assuming that additional sensing will always improve performance.

METHOD

Nineteen participants from an aerospace engineering programme were recruited. Their mean age was 26.2 years (SD = 3.37; range = 21–32 years), including 14 male and five female participants. All participants completed a simulated ATC task rather than operational air traffic control work. The experiment was approved by the Institutional Review Board of Nanyang Technological University under reference number IRB-2023-776. Five participants were excluded because their physiological data did not meet the required quality criteria. The final analysis therefore included 2,536 valid trials collected from 14 participants.

The simulated ATC task was presented using E-Prime 3.0 on a 23-inch desktop monitor with a resolution of 2560 × 1440 pixels (Figure 1). Eye movements were captured using a Tobii Pro Nano screen-based eye tracker.

EEG signals were recorded using an Emotiv EPOC X wireless headset. The simulation interface adopted aircraft symbols and labels similar to those used in an ATC research simulator. A section of the Singapore Flight Information Region was used as the operational background. Aircraft were presented together with their recent positions, call signs, flight levels, and speeds.



Figure 1: Physical setup of experiment and simulation interface.

Participants monitored multiple aircraft and determined whether conflicts were present. An AI conflict-resolution advisor assisted by identifying potential conflicts and recommending an altitude adjustment (Figure 2). When a conflict was detected, the interface displayed a red conflict warning, highlighted the relevant aircraft, presented distance rings, and provided a suggested altitude. When no conflict was detected, the advisor displayed a green no-conflict message.



(a) **(b)**

Figure 2: AI tool with no conflicts detected (a) and AI tool detecting conflict (b).

Each trial contained a monitoring phase and a decision phase. The monitoring phase began with a one-second fixation cross followed by a simulated flight sequence. At the beginning of the decision phase, the AI presented its assessment and, where applicable, an altitude recommendation. Participants combined the AI recommendation with their own judgement and entered a conflict-resolution response within the available time budget. They then rated their trust in the AI for that trial using a ten-point Likert scale. The experiment consisted of nine sections combining three AI reliability levels and three time budgets. The AI malfunctioned in 40% of trials under the low-reliability condition, 10% under the high-reliability condition, and

0% under the perfect-reliability condition. The time budgets were 6, 12, and 18 seconds. Each section contained 30 trials, and a two-minute break was provided between sections.

Eye-tracking data were recorded at 60 Hz and resampled to 50 Hz. Pupil diameter was calculated as the average of the left and right pupils and normalised using participant-level Z-scores. Fixations and saccades were identified using a velocity-threshold identification algorithm (Salvucci and Goldberg, 2000). Participant-specific velocity thresholds were estimated using an iterative procedure based on the distribution of gaze velocities (Nyström and Holmqvist, 2010). Extracted eye-tracking features included fixation count and duration, saccade count and duration, saccade amplitude, gaze coordinates, gaze velocity, and pupil-diameter statistics. Features were calculated separately for the monitoring phase, decision phase, and whole trial. Continuous EEG signals were acquired at 128 Hz and band-pass filtered between 1 and 50 Hz to reduce low-frequency drift and high-frequency noise while retaining the frequency range of interest. A 50-Hz notch filter was additionally applied to suppress power-line interference. Independent component analysis was used to identify and remove ocular artifacts, such as eye blinks. The signals remained at 128 Hz for subsequent processing. Time–frequency representations were computed using Morlet wavelets, and band-power measures were aggregated across the delta, theta, alpha, beta, and gamma frequency bands. The extracted feature set included mean, standard deviation, minimum, maximum, range, skewness, kurtosis, coefficient of variation, cross-band ratios, and temporal trends.

Trial-level trust ratings were divided into two classes. Ratings of 1–6 were labelled as low trust, while ratings of 7–10 were labelled as high trust. This threshold separated neutral-to-low ratings from clearly positive trust ratings. Five classification approaches were evaluated: Logistic Regression; Support Vector Machine; Random Forest; XGBoost; and one-dimensional Convolutional Neural Network. Models were trained separately using eye-tracking features, EEG features, and multimodal features. Classification was also performed separately for the monitoring phase, decision phase, and whole trial. GroupKFold cross-validation was used to ensure that data from the same participant were not simultaneously included in the training and test sets. The F1 score was used as the primary evaluation metric because it balances precision and recall.

RESULTS

After data-quality screening, 2,536 valid trials from 14 participants were retained, including 888 trials in the low-reliability condition, 775 in the high-reliability condition, and 873 perfect-reliability trials. Overall task accuracy was 93.8%. Mean trust increased with AI reliability, from 4.81 ± 2.55 under low reliability to 6.81 ± 1.88 under high reliability and 8.38 ± 0.83 under perfect reliability. Random Forest achieved the strongest performance, with whole-trial eye-tracking features producing the highest F1 score of 0.759 ± 0.059. Eye-tracking F1 scores were 0.740 ± 0.046 during monitoring and 0.757 ± 0.062 during decision-making (Figure 3).

As shown in Table 1, EEG-only performance was consistently lower, with F1 scores ranging from 0.646 to 0.691. Multimodal models achieved F1 scores between 0.680 and 0.703 and did not outperform the eye-tracking-only models. The results suggest that eye movements provided the most robust information for distinguishing trust states in this task. Adding EEG features did not automatically improve classification, despite previous findings suggesting advantages for multimodal sensing (Jiang et al., 2026).

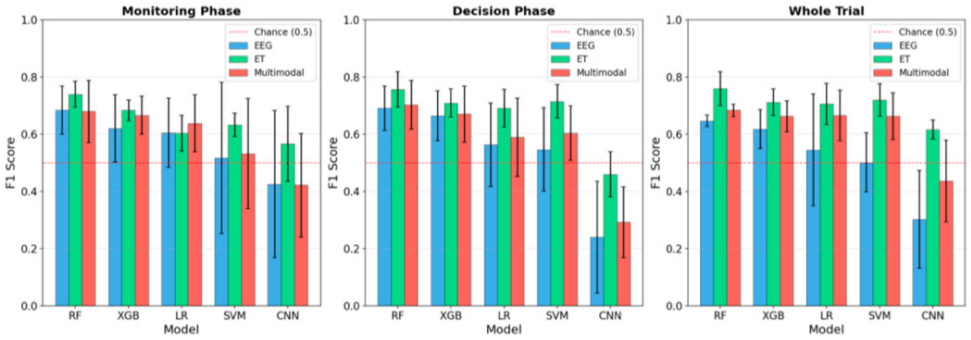


Figure 3: Bar graphs comparing model performances across phases and modalities.

Table 1: Summary of the best-performing models.

Phase	Modality	Model	F1 Score
Monitoring	Eye tracking	Random Forest	0.740 ± 0.046
Monitoring	EEG	Random Forest	0.685 ± 0.084
Monitoring	Multimodal	Random Forest	0.680 ± 0.109
Decision	Eye tracking	Random Forest	0.757 ± 0.062
Decision	EEG	Random Forest	0.691 ± 0.077
Decision	Multimodal	Random Forest	0.703 ± 0.086
Whole trial	Eye tracking	Random Forest	0.759 ± 0.059
Whole trial	EEG	Random Forest	0.646 ± 0.020
Whole trial	Multimodal	Random Forest	0.684 ± 0.021

Trust-Related Physiological Patterns: Pupil diameter was generally greater during low-trust trials (Figure 4). For whole-trial analysis, the mean normalised pupil diameter was −0.0515 ± 1.0122 during high trust and 0.1041 ± 0.9471 during low trust. Total fixation duration was shorter in high-trust trials. Participants accumulated an average total fixation duration of 7.90 ± 1.79 seconds under high trust, compared with 9.04 ± 2.57 seconds under low trust. This pattern is consistent with the interpretation that participants monitored and verified information for longer when they had less confidence in the AI recommendation. Longer fixation duration may reflect increased cognitive effort and more intensive information processing (Bruder and Hasse, 2020; Lu and Sarter, 2019).

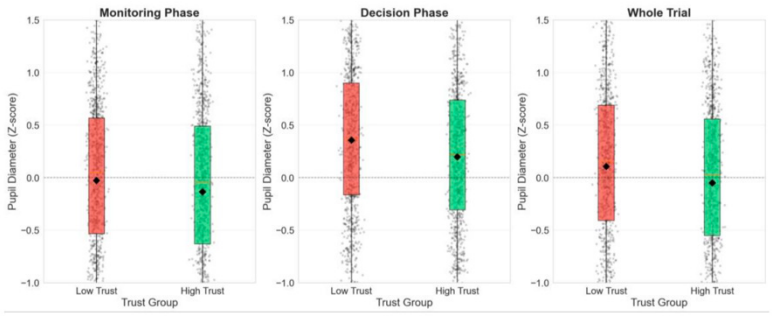


Figure 4: Boxplots of average pupil diameter.

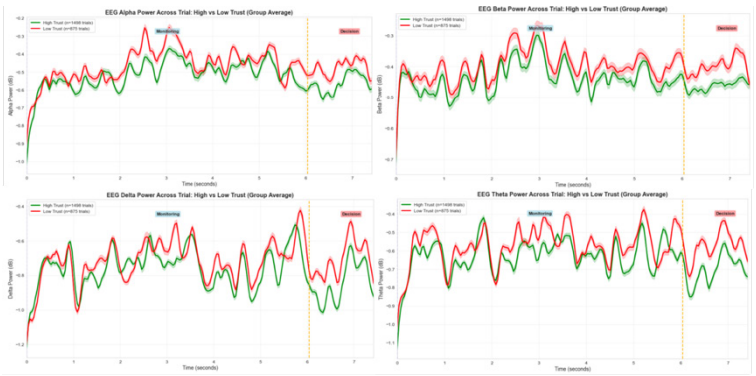


Figure 5: Average band-power trajectory across trials.

Average EEG band power was generally higher during low-trust trials than during high-trust trials (Figure 5). Beta-band measures were particularly relevant. Beta activity showed larger fluctuations and more extreme values during low-trust trials. Beta activity has previously been related to focused attention, cognitive readiness, stress, and active information processing (Attar, 2022). In the present context, lower trust may have required participants to maintain greater vigilance and evaluate the AI recommendation more critically.

DISCUSSION

The primary finding is that physiological signals can provide useful information for estimating trust during AI-assisted ATC monitoring. Random Forest models classified low and high trust with moderate performance, and eye-tracking features consistently produced stronger results than EEG or multimodal input. The best model achieved an F1 score of approximately 0.76 using whole-trial eye-tracking features.

The eye-tracking results are consistent with previous evidence that operator trust influences visual monitoring behaviour. Participants displayed longer fixation durations when trust was low, suggesting that they spent more time checking information or validating the AI recommendation. Previous studies

similarly found that lower trust or greater failure-detection demands can increase fixation duration and monitoring frequency (Hergeth et al., 2016; Körber et al., 2018). Greater pupil dilation during low-trust trials may further reflect increased risk perception, cognitive involvement, or workload (Lukovics et al., 2024).

Low-trust trials were also characterised by higher and less stable EEG activity, particularly in the beta band. This finding may indicate heightened cognitive readiness and information processing when participants were uncertain about AI performance. However, EEG findings should be interpreted cautiously. The preprocessing procedure averaged band power across EEG channels, removing spatial information that may have been relevant to classification. This may partly explain why EEG-only models underperformed relative to eye tracking.

The lack of improvement from multimodal input is an important practical result. Multimodal sensing is often expected to outperform unimodal sensing because different signals may contain complementary information (Jiang et al., 2026). In this study, however, adding EEG features did not improve performance beyond eye tracking alone. This may reflect noise in the EEG signals, limited sample size, the loss of channel-level spatial information, or an ineffective fusion strategy. From an applied perspective, the result suggests that a simpler eye-tracking system may offer a better balance between predictive value, wearability, and operational burden in future trust-monitoring applications.

The study contributes to human-factors research in three ways. First, it provides physiological evidence of trust responses in a simulated AI-assisted ATC task, a context that remains less studied than automated driving. Second, it compares eye tracking, EEG, and multimodal sensing under the same experimental conditions. Third, it identifies interpretable physiological patterns that may support the future development of adaptive human–AI interfaces.

CONCLUSION

This study investigated physiological indicators of trust during an AI-assisted air traffic control simulation. Eye-tracking and EEG features were used to classify low and high subjective trust across monitoring, decision, and whole-trial phases. Random Forest achieved the strongest results, while eye tracking consistently outperformed EEG and multimodal input. Low trust was associated with longer fixation duration, greater pupil dilation, and higher or less stable EEG activity, particularly within the beta band. These findings demonstrate the potential of physiological sensing for monitoring trust in safety-critical human–AI systems. The results also highlight the importance of practical sensor selection. More modalities do not necessarily produce better trust recognition. In the present task, interpretable eye-tracking measures provided the strongest predictive performance with lower sensing complexity. Future work should validate these findings with professional air traffic controllers, retain spatial EEG information, and develop adaptive interfaces capable of supporting trust calibration without increasing operator workload.

ACKNOWLEDGMENT

This work was supported by the Fundamental Research Funds for the Central Universities under Grant No. 3120626016-104972026RSCbs0015 and by the Hubei Postdoctoral Advanced Talent Program under Grant No. 2025HBBSHJD031.

REFERENCES

Akash, K., Hu, W.-L., Jain, N., Reid, T., 2018. A classification model for sensing human trust in machines using EEG and GSR. ACM Trans Interact Intell Syst 8, 27:1-27:20. https://doi.org/10.1145/3132743

Attar, E.T., 2022. Review of electroencephalography signals approaches for mental stress assessment. Neurosci. J. 27, 209–215. https://doi.org/10.17712/nsj.2022.4.20220025

Bestugin A.R., Eshenko A.A., Filin A.D., Plyasovskikh A.P., Shatrakov A.Y., Shatrakov Y.G., 2020. Air Traffic Control Automated Systems, Springer Aerospace Technology. Springer, Singapore. https://doi.org/10.1007/978-981-13-9386-0

Bruder, C., Hasse, C., 2020. What the eyes reveal: Investigating the detection of automation failures. Appl. Ergon. 82, 102967. https://doi.org/10.1016/j.apergo.2019.102967

Chiou, E.K., Lee, J.D., 2023. Trusting Automation: Designing for Responsivity and Resilience. Hum. Factors J. Hum. Factors Ergon. Soc. 65, 137–165. https://doi.org/10.1177/00187208211009995

Clark, R.M., Ingebritsen, A.M., Bustamante, E.A., 2010. Differential Effects of Likelihood Alarm Technology and False-Alarm vs. Miss-Prone Automation on Decision-Making Accuracy and Bias. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 54(19), 1508–1512. https://doi.org/10.1177/154193121005401932.

Hergeth, S., Lorenz, L., Vilimek, R., Krems, J.F., 2016. Keep Your Scanners Peeled: Gaze Behavior as a Measure of Automation Trust During Highly Automated Driving. Hum. Factors J. Hum. Factors Ergon. Soc. 58, 509–519. https://doi.org/10.1177/0018720815625744

Hoff, K.A., Bashir, M., 2015. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. Hum. Factors J. Hum. Factors Ergon. Soc. 57, 407–434. https://doi.org/10.1177/0018720814547570

Jiang, C., Zhou, T., Shen, H., Wang, Y., 2026. Understanding Trust and Fatigue Under Transparency Regulation in Air Traffic Control: A Multimodal Approach. IEEE Intell. Transp. Syst. Mag. 18, 21–35. https://doi.org/10.1109/MITS.2025.3636860

Kohn, S.C., De Visser, E.J., Wiese, E., Lee, Y.-C., Shaw, T.H., 2021. Measurement of trust in automation: A narrative review and reference guide. Front. Psychol. 12, 604977. https://doi.org/10.3389/fpsyg.2021.604977

Körber, M., Baseler, E., Bengler, K., 2018. Introduction matters: Manipulating trust in automation and reliance in automated driving. Appl. Ergon. 66, 18–31. https://doi.org/10.1016/j.apergo.2017.07.006

Li, Q., Ng, K.K.H., Yiu, C.Y., Yuan, X., So, C.K., Ho, C.C., 2023. Securing air transportation safety through identifying pilot's risky VFR flying behaviours: An EEG-based neurophysiological modelling using machine learning algorithms. Reliab. Eng. Syst. Saf. 238, 109449. https://doi.org/10.1016/j.ress.2023.109449

Li, Z., Li, F., Xu, G., Li, D., 2025. Beyond the Gaze: Peripheral Vision-Aware Visual Detection Failures Recognition Through LLM-Based Fixation Coordinate-Sensitive Analysis. IEEE Trans. Intell. Transp. Syst. 26, 20766–20785. https://doi.org/10.1109/TITS.2025.3588621

Liu, B., Lye, S.W., Yeo, K.X., Chen, C.-H., 2025. A human-centric model for task demand assessment based on unsupervised learning-assisted eye movement measure. Adv. Eng. Inform. 65, 103259. https://doi.org/10.1016/j.aei.2025.103259

Liu, B., Lye, S.W., Zakaria, Z.B., 2024. An integrated framework for eye tracking-assisted task capability recognition of air traffic controllers with machine learning. Adv. Eng. Inform. 62, 102784. https://doi.org/10.1016/j.aei.2024.102784

Liu, Q., Li, F., Ng, K.K.H., Han, J., Feng, S., 2025. Accident investigation via LLMs reasoning: HFACS-guided Chain-of-Thoughts enhance general aviation safety. Expert Syst. Appl. 269, 126422. https://doi.org/10.1016/j.eswa.2025.126422

Lu, Y., Sarter, N., 2019. Eye Tracking: A Process-Oriented Method for Inferring Trust in Automation as a Function of Priming and System Reliability. IEEE Trans. Hum.-Mach. Syst. 49, 560–568. https://doi.org/10.1109/THMS.2019.2930980

Lukovics, M., Prónay, S., Nagy, B., 2024. Segmented Trust Assessment in Autonomous Vehicles via Eye-Tracking. J. Intell. Connect. Veh. 7, 151–161. https://doi.org/10.26599/JICV.2023.9210037

Lyu, M., Li, F., 2025. Do you need help? Identifying and responding to pilots' trouble shooting through eye-tracking and large language model. Int. J. Hum.-Comput. Stud. 205, 103617. https://doi.org/10.1016/j.ijhcs.2025.103617

Lyu, M., Li, Q., 2026. Methodological Challenges in Human Factors Research for Air Traffic Management, in: 2026 IEEE International Conference on Human-Machine Systems (ICHMS). Presented at the 2026 IEEE International Conference on Human-Machine Systems (ICHMS), pp. 71–76. https://doi.org/10.1109/ICHMS69701.2026.11602298

Nyström, M., Holmqvist, K., 2010. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. Behav. Res. Methods 42, 188–204. https://doi.org/10.3758/BRM.42.1.188

Oh, S., Seong, Y., Yi, S., Park, S., 2020. Neurological Measurement of Human Trust in Automation Using Electroencephalogram. Int. J. FUZZY Log. Intell. Syst. 20, 261–271. https://doi.org/10.5391/IJFIS.2020.20.4.261

Salvucci, D.D., Goldberg, J.H., 2000. Identifying fixations and saccades in eye-tracking protocols, in: Proceedings of the 2000 Symposium on Eye Tracking Research & Applications, ETRA '00. Association for Computing Machinery, New York, NY, USA, pp. 71–78. https://doi.org/10.1145/355017.355028

Sarker, A., Shen, H., Rahman, M., Chowdhury, M., Dey, K., Li, F., Wang, Y., Narman, H.S., 2020. A Review of Sensing and Communication, Human Factors, and Controller Aspects for Information-Aware Connected and Automated Vehicles. IEEE Trans. Intell. Transp. Syst. 21, 7–29. https://doi.org/10.1109/TITS.2019.2892399

Walker, F., Wang, J., Martens, M.H., Verwey, W.B., 2019. Gaze behaviour and electrodermal activity: Objective measures of drivers' trust in automated vehicles. Transp. Res. Part F Traffic Psychol. Behav. 64, 401–412. https://doi.org/10.1016/j.trf.2019.05.021

Xia, Z., Chen, C.-H., Hsieh, M.-H., Lim, J.D., Sng, H.C.K., Ahmad, M.K., 2025a. Detection of in-the-wild and long-term stress based on cross-attention transformer and context-aware ensemble model. Adv. Eng. Inform. 68, 103586. https://doi.org/10.1016/j.aei.2025.103586

Xia, Z., Chen, C.-H., Hsieh, M.-H., Ma, G., Liu, B., Yu, X., Xin, S., Dong, L., 2026. A systematic review on human-AI hybrid systems and human factors in air traffic management. J. Eng. Des. 37, 871–919. https://doi.org/10.1080/09544828.2025.2509056

Xia, Z., Chen, C.-H., Kuo, J.-Y., Zhang, M., 2025b. How Do Stress Situations Affect Higher-Level Text Processing in L1 and L2 Readers? An Eye-Tracking Study. J. Eye Mov. Res. 18, 7. https://doi.org/10.3390/jemr18020007

Yao, X., Chen, C.-H., Liu, B., Ma, G., Yu, X., 2026a. An explainable eye-tracking-based framework for enhanced level-specific situational awareness recognition in air traffic control. Adv. Eng. Inform. 69, 103928. https://doi.org/10.1016/j.aei.2025.103928

Yao, X., Chen, C.-H., Liu, B., Zhou, Y., Yu, X., 2026b. Toward Safer Human–AI Teaming: Phase-Specific Recognition and Interpretation of Automation Trust in Air Traffic Control. Int. J. Human–Computer Interact. 1–24. https://doi.org/10.1080/10447318.2026.2689527

Yao, X., Jing, C., Huang, Y., Zhi, J., 2022. A comparative study of passenger multitasking activities on commuting and leisure electrified intercity railways. J. Adv. Transp. 2022, 3001392. https://doi.org/10.1155/2022/3001392

Yu, D., Tang, X., Xie, X., Zhi, J., 2026. LitePCT-BR: An interpretable multi-scale architecture for the recognition and prediction of train driver attentional decline. Adv. Eng. Inform. 73, 104585. https://doi.org/10.1016/j.aei.2026.104585

Yu, D., Wang, W., Ding, T., Niu, Z., Zhi, J., Xie, X., Zhang, Z., Tang, X., Yu, Dongyu, 2025. An application study of a DynGraph-TF-based EEG recognition framework for predicting mental fatigue in train drivers. Adv. Eng. Inform. 68, 103675. https://doi.org/10.1016/j.aei.2025.103675

Yu, X., Chen, C.-H., Yang, H., 2023. Air traffic controllers' mental fatigue recognition: A multi-sensor information fusion-based deep learning approach. Adv. Eng. Inform. 57, 102123. https://doi.org/10.1016/j.aei.2023.102123

Yu, X., Yang, H., Chen, C.-H., 2025. Human operators' cognitive workload recognition with a dual attention-enabled multimodal fusion framework. Expert Syst. Appl. 280, 127418. https://doi.org/10.1016/j.eswa.2025.127418

Zhang, T., Yang, J., Chen, M., Li, Z., Zang, J., Qu, X., 2024. EEG-based assessment of driver trust in automated vehicles. Expert Syst. Appl. 246, 123196. https://doi.org/10.1016/j.eswa.2024.123196