Applied Human Factors and Ergonomics (AHFE2025), Vol. 199, 2025, 2016–2021
https://doi.org/10.54941/ahfe1007013

Tackling Human Factors in Aviation Safety - An Application of AI Facial Recognition Technology

Yanbing Chen[1], Shuzhen Luo[2], and Andy Dattel[1]

[1]School of Graduate Studies, College of Aviation, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA
[2]Mechanical Engineering, College of Engineering, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA

ABSTRACT

Human factors remain a predominant cause of transportation accidents with physiological, psychological, and emotional states significantly influencing operator performance. In aviation, factors such as fatigue, stress, illness, medication, and substance use impair pilot performance, leading to compromised decision-making, reduced situational awareness, and increased risk-taking behavior. While regulatory guidelines and medical evaluations exist to address these challenges, current measures often rely on self-reporting and subjective assessments that can be prone to bias. This research proposes an Artificial Intelligence (AI)-driven facial recognition model to objectively assess pilot fitness to fly by analyzing micro-expressions (eye corners and mouth edges—to measure the eye aspect ratio, mouth aspect ratio, assisted by the percentage of eye closure). The study is structured into three sequential tasks: 1) model development using publicly available facial image datasets under varied physiological and emotional conditions; 2) laboratory-based validation involving healthy participants performing cognitive tasks under stress to refine detection accuracy; and 3) aviation-specific validation using flight simulators with licensed pilots performing tasks under normal and high-stress conditions. Data from facial imagery, simulator performance metrics, and physiological measures (e.g., heart rate variability) will be integrated to enhance model precision. Deep convolutional neural networks with transfer learning will classify facial features linked to fatigue, stress, illness, or substance influence. This research aims to establish a non-intrusive, real-time AI system capable of improving aviation safety by providing objective assessments of pilot fitness, potentially extending to other transportation domains.

Keywords: Fatigue, Machine learning, Risk management, Artificial intelligence, Pilot safety

INTRODUCTION

Transportation accidents are often attributed to human factors, which account for a significant proportion of incidents and fatalities. Reports from organizations such as the National Transportation Safety Board (Marcus and Rosekind, 2017) and the Federal Aviation Administration (FAA) consistently identify human error as a leading cause of accidents (FAA, 2006). These errors are frequently influenced by physiological, psychological, and emotional factors, underscoring the critical need to address human limitations to improve safety outcomes. Beyond aviation, similar issues have been observed in rail, maritime, and road transportation, where human performance plays a central role in preventing accidents and ensuring operational efficiency.

Human factors, including illness, medication, alcohol consumption, fatigue, stress, and emotional state, play critical roles in transportation safety (Alavi et al., 2017; Shandhana Rashmi and Marisamynathan, 2023). Even minor illnesses can severely impair pilot performance by reducing judgment, alertness, memory, and coordination. While medications may control symptoms, they can introduce side effects that degrade cognitive and motor abilities (Hafez et al., 2023). For instance, sedatives, antihistamines, and other medications may impair alertness and heighten vulnerability to hypoxia. The Federal Aviation Regulations explicitly prohibit flying under the influence of any substance that compromises safety, emphasizing the need for professional medical guidance when assessing fitness to fly.

Fatigue is a significant, often hidden risk in commercial aviation. Currently, crew members self-assess their fitness for duty at sign-in (Jackson and Earl, 2006; Tvaryanas and MacPherson, 2009); however, fatigue is a contributing factor in 15–20% of transportation accidents (Akerstedt, 2000; Jun-Ya and Rui-Shan, 2023). Regulators set duty and flight time limits to reduce the risk of fatigue, but these do not address all factors affecting pilot performance, partly due to the missed physiological alertness cues. Aviation requires objective, real-time, non-intrusive fitness tools to address this gap.

Artificial Intelligence (AI) driven facial recognition model has been used in other industries to assess human subjects' health status (Chan et al., 2024) and cognitive workload (Iarlori et al., 2024). This research aims to develop an AI-driven facial recognition model to objectively assess pilot fitness to fly by analyzing micro expressions, facial symmetry, eye movement, and other biomarkers that reflect fatigue, stress, and impairment. The AI model will be trained using publicly available datasets containing facial images of individuals in varying conditions such as fatigue, drowsiness, stress, sadness, and under the influence of alcohol, drugs, or medication. Data preprocessing will employ facial landmark detection and attention-based image segmentation to isolate key facial regions, including the eyes (tracking movement and redness), mouth (symmetry, dryness, or tremor), and skin tone (color changes indicative of intoxication or stress) (Chan et al., 2024). Model training will leverage deep convolutional neural networks, utilizing transfer learning techniques to enhance performance with smaller datasets.

METHODOLOGY

There are three tasks in this research. Task 1 focuses on model building using secondary data from publicly available facial image datasets in different conditions. Task 2 involves a laboratory-based experiment with healthy individuals to validate and refine the AI algorithm's accuracy in detecting cognitive performance changes under stress. Participants (n = 12) will perform cognitive tasks under high-stress conditions, and facial images will be captured to fine-tune the algorithm. Task 3 includes a pilot simulation-based experiment to fine-tune the AI algorithm for aviation-specific applications. Licensed pilots (n = 12) will perform flight simulation tasks under high-workload or stressful conditions, such as emergency scenarios and adverse weather conditions. Data from facial images and simulator metrics like decision-making speed, navigation accuracy, and task prioritization will be analyzed to adapt the AI algorithm for real-time, aviation-specific assessments. The study chronology is presented as Figure 1 as a closed-loop validation design.

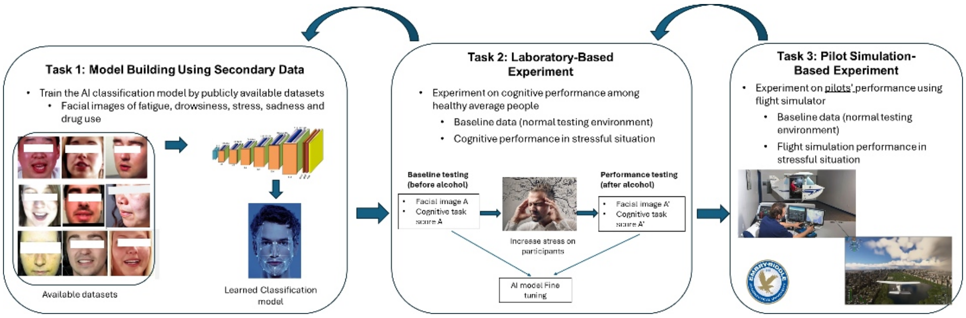

Figure 1: Study chronology.

Procedures

Task 1 aims to train an AI classification model using publicly available facial images of individuals exhibiting signs of fatigue, stress, emotional distress, illness, or under the influence of alcohol, drugs, or medication. Publicly available datasets containing facial images of individuals in various conditions will be collected to allow the AI model to focus on relevant visual cues. For example:

- Fatigue and drowsiness: Datasets with images of individuals across various states of alertness will be sourced.
- Stress and sadness: Emotional facial expression datasets will be used.
- Alcohol, drugs, and medication: Datasets with varying levels of sobriety and intoxication will be included, such as sober, mildly intoxicated, and heavily intoxicated conditions.

Data analysis: Facial images will undergo preprocessing using techniques like facial landmark detection and attention-based image segmentation to isolate key facial regions, including:

- Eyes (tracking movement and redness)
- Mouth (symmetry, dryness, or tremor)
- Skin tone (color changes indicative of intoxication or stress).

Model Training: A deep convolutional neural network will be developed to classify facial images and identify relationships between facial features and specific states (e.g., fatigue and stress). Transfer learning techniques may be used to leverage pre-trained models for improved performance with smaller datasets (Alshardan et al., 2024).

Task 2 aims to validate the initial AI Facial Recognition algorithm by assessing its ability to detect performance changes among individuals under controlled conditions in the lab. First, participants (n = 12), not necessarily pilots, will perform cognitive tasks focusing on deductive reasoning, working memory, concentration, associative learning, and spatial planning. Second, participants will repeat the same cognitive tasks under high-stress scenarios, such as time-constrained problem-solving or multitasking under simulated pressure. Their facial images and performance data will be used to refine the algorithm. The heart rate variability will also be monitored as another measurement to examine AI model accuracy.

Task 3 aims to fine-tune the AI algorithm to detect performance change specific to pilots working under stressful or high-workload environments. First, licensed pilots (n = 12) Participants will perform flight tasks on a simulator under normal conditions to establish baseline performance. Second, pilots will perform the same simulation tasks when exposed to stressful or high-workload conditions. For example, pilots will undertake challenging flight scenarios, such as emergency procedures and simulated adverse weather conditions. Facial images of pilots will be captured during baseline and high-stress tasks using high resolution webcam in front of them. Data from simulator metrics, such as decision-making speed, navigation accuracy, and task prioritization, will be analyzed after the experiment. The heart rate of each participant is also monitored throughout the experiment process.

Data Analysis

There are three types of data to be analyzed in this study: facial image data, performance data, and physiological data. Facial image data collected in Task 1 are publicly available data used to train AI to build a basic model to recognize people under influence of fatigue, stress, illness, medication, and substance use etc. Facial images collected in Task 2 & 3 will be used to validate the basic AI model together with the statistical analysis results to improve the classification accuracy of the model. These findings can validate and fine-tune the AI model by ensuring that facial recognition technology can accurately detect stress-induced features, such as skin color and mouth symmetry. Physiological data such as heart rate and stress level to be collected in Task 2 and Task 3 will be used to validate the research outcome. Pearson Correlation will be used to evaluate the relationship between heart rate variability and cognitive/flight performance.

SIGNIFICANCE

This research is significant as it addresses the critical role of human factors including emotional, psychological, and physiological influences on transportation safety. By examining factors such as stress, fatigue, cognitive overload, and substance use impact operator performance, this study provides a rapid approach to predict human factor related risks in aviation, and the use of which has a potential to be extended to other types of transportation including maritime, rail, and road to improve safety. To date, real-time eye-trackers are frequently used to predict driver fatigue through participants' eye-movement data. However, false alarms often occur when participants wear sunglasses. Such a problem may be able to be solved by using AI facial recognition technology, as facial characteristics collected are from the whole face to make the predication.

CONCLUSION

The findings from this research are also essential for developing targeted interventions to improve safety outcomes. For example, identifying the effects of emotional states and psychological pressures on decision-making can inform training programs that build emotional resilience and enhance situational awareness. The research outcome will assist with the iteration of robust policies, monitoring systems, and workplace wellness initiatives to mitigate these risks. Additionally, this AI-assisted approach contributes to the broader goal of enhancing transportation safety by informing policy, advancing training methodologies, and fostering a proactive safety culture. The findings will benefit transportation operators, industry stakeholders, policymakers, and safety regulators by offering evidence-based strategies to minimize risks, improve performance, and ensure the safety and efficiency of transportation systems nationally.

ACKNOWLEDGMENT

The authors would like to acknowledge Embry-Riddle Aeronautical University FIRST Program to fund the research. We would like to thank the School of Graduate Studies in College of aviation and Mechanical Engineering department in College of Engineering for the consistent support.

REFERENCES

Akerstedt, T. (2000) "Consensus statement: Fatigue and accidents in transport operations," *Journal of Sleep Research*, 9(4), p. 395. Available at: https://doi.org/10.1046/J.1365–2869.2000.00228. X.

Alavi, S. S. et al. (2017) "Personality, Driving Behavior and Mental Disorders Factors as Predictors of Road Traffic Accidents Based on Logistic Regression," Iranian Journal of Medical Sciences, 42(1), p. 24. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC5337762/ (Accessed: January 19, 2025).

Alshardan, A. et al. (2024) "Leveraging transfer learning-driven convolutional neural network-based semantic segmentation model for medical image analysis using MRI images," Scientific Reports 2024 14:1, 14(1), pp. 1–26. Available at: https://doi.org/10.1038/s41598–024-81966-y.

Chan, J. J. H. et al. (2024) "Facial artificial intelligence in ophthalmology and medicine: Fundamental and transformative applications," Therapeutic Advances in Ophthalmology, 16, p. 25158414241302870. Available at: https://doi.org/10.1177/25158414241302871.

FAA (2006) Human Factors Guidance for Intelligent Transportation Systems at the Highway-Rail Federal Railroad Administration Intersection.

Hafez, G. et al. (2023) "Drugs with a negative impact on cognitive functions (Part 2): drug classes to consider while prescribing in CKD patients," Clinical Kidney Journal, 16(12), p. 2378. Available at: https://doi.org/10.1093/CKJ/SFAD239.

Iarlori, S. et al. (2024) "An Overview of Approaches and Methods for the Cognitive Workload Estimation in Human–Machine Interaction Scenarios through Wearables Sensors," BioMedInformatics 2024, 4(2), pp. 1155–1173. Available at: https://doi.org/10.3390/BIOMEDINFORMATICS4020064.

Jackson, C. A. and Earl, L. (2006) "Prevalence of fatigue among commercial pilots," Occupational Medicine, 56(4), pp. 263–268. Available at: https://doi.org/10.1093/OCCMED/KQL021.

Jun-Ya, S. and Rui-Shan, S. (2023) "Pilot fatigue survey: A study of the mutual influence among fatigue factors in the 'work' dimension," Frontiers in Public Health, 11, p. 1014503. Available at: https://doi.org/10.3389/FPUBH.2023.1014503/BIBTEX.

Marcus, J. H. and Rosekind, M. R. (2017) "Fatigue in transportation: NTSB investigations and safety recommendations," Injury Prevention, 23(4), pp. 232–238. Available at: https://doi.org/10.1136/INJURYPREV-2015–041791.

Shandhana Rashmi, B. and Marisamynathan, S. (2023) "Factors affecting truck driver behavior on a road safety context: A critical systematic review of the evidence," Journal of Traffic and Transportation Engineering (English Edition), 10(5), pp. 835–865. Available at: https://doi.org/10.1016/J. JTTE.2023.04.006.

Tvaryanas, A. P. and MacPherson, G. D. (2009) "Fatigue in pilots of remotely piloted aircraft before and after shift work adjustment," Aviation Space and Environmental Medicine, 80(5 PART 1), pp. 454–461. Available at: https://doi.org/10.3357/ASEM.2455.2009.