



ARL-TR-10048 • JAN 2025



Final Report on Multimodal Inference of Human State: Tracking Cognition in a Risky Environment

by Jason S. Metcalfe, W. David Hairston, Namazbai Ishmakhametov, Mohammad Y. M. Naser, and Sylvia Bhattacharya

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.



Final Report on Multimodal Inference of Human State: Tracking Cognition in a Risky Environment

Jason S. Metcalfe and W. David Hairston

DEVCOM Army Research Laboratory

Namazbai Ishmakhametov, Mohammad Y. M. Naser, and Sylvia Bhattacharya

Neuro-Interaction Innovation Laboratory, Kennesaw State University

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
January 2025	Technical Report	START DATE 8/26/2020	END DATE 8/26/2024
4. TITLE AND SUBTITLE Final Report on Multimodal Inference of Human State: Tracking Cognition in a Risky Environment			
5a. CONTRACT NUMBER W911NF2020205		5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
5d. PROJECT NUMBER		5e. TASK NUMBER	5f. WORK UNIT NUMBER
6. AUTHOR(S) Jason S. Metcalfe, W. David Hairston, Namazbai Ishmakhametov, Mohammad Y. M. Naser, and Sylvia Bhattacharya			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) DEVCOM Army Research Laboratory ATTN: FCDD-RLA-FB Aberdeen Proving Ground, MD 21005			8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-10048
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) US ARMY ACC-APG-RTP W91		10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.			
13. SUPPLEMENTARY NOTES ORCIDs: Jason S. Metcalfe, 0000-0001-9086-9962; W. David Hairston, 0000-0003-4432-8430			
14. ABSTRACT This report details the execution and outcomes of a 3-year collaboration between the U.S. Army Combat Capabilities Development Command Army Research Laboratory and Kennesaw State University that focused multimodal data integration to improve human state inference in complex, high-risk environments. The project advanced signal processing techniques and machine learning algorithms for classifying human states related to ongoing, task-related events. Accompanied by seven publications, data, and a code repository, the primary contribution from this program was the Sequential, Event-Based Multimodal Fusion framework, which enhances the integration of diverse physiological, behavioral, and environmental data streams. This framework improved accuracy of classifying events from DEVCOM Army Research Laboratory within real-world driving experiments. These advancements support ARL's mission toward human-machine integration for future operational formations. The project also contributed to talent development by involving 28 students across various disciplines, providing a cross-functional team of students with hands-on experience in multimodal data fusion and Army-focused research. This engagement supported career transitions into industry and academia, strengthening the national workforce that will manifest future Army technology-based advantages. This project further established an interdisciplinary research ecosystem, and the outcomes advance ARL's objectives in the broader scientific space of multimodal machine learning and classification.			
15. SUBJECT TERMS Humans in Complex Systems, machine learning, multimodal, human state inference, wearable sensors, autonomy, driving automation, automobile passengers, real-world driving, open road driving, EEG, electroencephalography, ECG, electrocardiography, psychophysiology			
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED	
19a. NAME OF RESPONSIBLE PERSON Jason S. Metcalfe			18. NUMBER OF PAGES 97
			19b. PHONE NUMBER (Include area code) (410) 278-9229

STANDARD FORM 298 (REV. 5/2020)

Prescribed by ANSI Std. Z39.18

Contents

List of Figures	v
List of Tables	vi
Preface	vii
Acknowledgments	ix
1. Introduction	1
1.1 Background	2
1.2 Army Research Context	2
2. Project Overview	6
2.1 Initial Program Plan	6
2.2 Program Plan Development	8
3. Execution, Evolution, and Results	9
3.1 Executed Program Overview	9
3.2 Overview of Outcomes	11
3.2.1 Toward Data-Driven Insights into Passenger Experiences	11
3.2.2 Future Multimodal System Concepts for Human–Machine Integrated Formations	16
3.2.3 A Revised Multimodal Vision	19
4. Key Highlights, Values, Insights, and Outcomes	20
4.1 Ecosystem Growth, Diversification, and Talent Cultivation	21
4.2 Basic and Applied Research Contributions	22
4.3 Evolving Future Concepts to Deliver Transformational Overmatch	24
5. Conclusion	25
6. References	26

Appendix A. About the Collaborators	30
Appendix B. Student Participation	35
Appendix C. Research Dissemination	42
Appendix D. Technical Supplement	64
List of Symbols, Abbreviations, and Acronyms	84
Distribution List	86

List of Figures

Figure 1.	RWD study timeline and data elements.	4
Figure 2.	Initial program plan for FY2021, with FY2022 and FY2023 options.	7
Figure 3.	Executed program, FY2021–FY2023.	10
Figure 4.	Average brain activity for a passenger-driver (upper-lower) dyad during three driving events.....	13
Figure 5.	Custom system to extract head pose for each video frame.	16
Figure 6.	Conceptual brain-based control system (based on Biswas et al. 2023).....	18
Figure 7.	Sequential Event-based Multimodal Fusion ([SEMF] developed in NIIL at KSU) yields ensemble outputs from hierarchical, multimodal feature-based decision layers.	19
Figure A-1.	Location of KSU campuses in the greater Atlanta, Georgia area. ..	31
Figure A-2.	Members of the Neuro-Interaction Innovation Laboratory in 2023.	32
Figure A-3.	Sylvia Bhattacharya, PhD.	33
Figure B-1.	Distribution of NIIL student major areas of study.	37
Figure B-2.	Map of places of origin of NIIL student members.....	37
Figure C-1.	Distribution of h-indices for all authors of published works cited in external papers in this appendix.	44
Figure C-2.	RWD study information.	46
Figure C-3.	Exemplar postprocessed EEG data recorded during the RWD study.	47
Figure C-4.	Instrumentation and configuration for head pose classification during driving.....	49
Figure C-5.	Suboptimal head pose classification system.	50
Figure C-6.	Results and accuracy of classification applied to 26 passengers.....	50
Figure C-7.	Results describing brain activity preceding driving events.....	52
Figure C-8.	Data from passengers and drivers, left and right columns, respectively, by driving event type (top to bottom: aggressive acceleration, braking, lane change).	53
Figure C-9.	Data from Figure C-8, decomposed across frequency within each time-step.	53
Figure C-10.	Illustration of behavioral limits in roadway vehicles under safety restraint (images generated with support of GPT 4o).	55
Figure C-11.	Comparison of several classification methods applied to smart seat sensors.	55

Figure C-12. Examples of common methods for observing and recording brain activity.	57
Figure C-13. Overview of traditionally applied BCI and brain-measurement methods.	58
Figure C-14. Fixed vs. flexible solution paths in complex goal spaces for decision agents.	60
Figure C-15. Summary of the proposed method for evaluation of the rigor or physiological ML research reports.	62
Figure C-16. Charts representing the reported use of relatively small samples that are imbalanced toward biologically male-heavy study populations.	63
Figure D-1. Illustration of the SEMF approach.	73
Figure D-2. Preliminary classification results for multiclass vs. sequential, binary methods.	77

List of Tables

Table 1. List of appendixes.	5
Table B-1. Student data by major area of study and degree level.	36
Table C-1. Statistics for journals, papers, and people who have cited the works in this appendix.	43
Table D-1. Technical deliverables: data, documentation, and software.	65
Table D-2. Summary of passenger highway experience.	66
Table D-3. Passenger data features for real-world event classification.	68
Table D-4. Classifier accuracies (%) for uni- and multimodal methods applied to binary problems.	75
Table D-5. Code catalog.	80

Preface

This statement pertains specifically to the use and application of outputs from intelligent text synthesis products in the generation of the contents herein. By intelligent text synthesis, we specifically refer to the application of technologies that are based on large language models (LLMs). The lead author, JSM, took sole responsibility for the appropriate leveraging of multiple LLM-based tools, otherwise known colloquially as “AI,” during the production of this report.

LLM and AI tools were used as summarizers and organizers of human-input information. LLMs and AI tools were never used as a primary content generator of any section or subsection of this document. Any prompting sessions resulting in text inputs that were successively integrated into this document began with document-based inputs written by one or more of the authors and/or student collaborators. Inputs shared with the LLM tools included laboratory meeting notes, graphic and tabular content as screen-captured during group laboratory meetings, all published works by this group, as well as generally focused, but loosely structured instruction sets that were actively iterated (and corrected for false outputs and “hallucinations” manually by the lead author doing the prompting); the lead author also paused and conducted research whenever conceptual or factual ambiguities required resolution.

As much as possible—with the limits of human error—non-public materials shared with the LLMs as input-prompt materials (e.g., laboratory notes) were carefully read, as well as searched using available text tools, for any sensitive contents (personal, personally identifiable information, or evolving technical details) that, when identified, were redacted before inclusion in the LLM prompt.

A first full-text version of the publicly releasable report (version 1.0) was also submitted—redacted for inclusions with specific persons and entities (i.e., U.S. Army Combat Capabilities Development Command, Army Research Laboratory, Army, Kennesaw State University, and Neuro-Interaction Innovation Laboratory), as well as all graphic content—for report consolidation before first external editorial review, as well as approval confirmation from all co-authors. The final version of this document sent for technical and scientific review was deemed completed by a final full human review for finalization, making any final adjustments and edits manually.

At all stages of integration, LLM inputs were solicited based on and constrained to a set of information derived from expert, or subject matter expert, human-authored input, subject to online and follow-up human review and correction, and were ultimately integrated into a human-designed and human-controlled master document by a human author. Interested readers may contact JSM (jason.s.metcalf2.civ@army.mil) for further discussion and details.

Acknowledgments

We gratefully acknowledge Dr. Joe Rexwinkle (U.S. Army Combat Capabilities Development Command Army Research Laboratory, Aberdeen Proving Ground [APG], Maryland) for reviewing the manuscript, providing feedback, and assisting with the Form 1 process. We also acknowledge Drs. Kristin E. Schaefer-Lay (DEVCOM Army Research Laboratory, APG, Maryland) and Jean M. Vettel (Army Futures Command Headquarters, Austin, Texas) for their efforts in reviewing the work and providing feedback. Finally, we are deeply grateful to Jean M. Vettel for her essential role as the co-Principal Investigator on the real-world driving experiment, which formed the basis of this entire cooperative agreement. All contributors provided suggestions that fundamentally steered the focus and clarity of this report.

1. Introduction

This report documents the formation, execution, and outcomes of the Cooperative Agreement between the U.S. Army Combat Capabilities Development Command Army Research Laboratory and Kennesaw State University (KSU). The Cooperative Agreement enabled a 3-year project entitled Multimodal Inference of Human State to Track Cognitive Processes in Risky Environments,* led by Dr. Sylvia Bhattacharya of the Neuro-Interaction Innovation Laboratory (NIIL) in the Southern Polytechnic College of Engineering & Engineering Technologies at the KSU campus in Marietta, Georgia. This program addressed a fundamental challenge for the Army regarding how to establish information security and maintain decision dominance; that is, how to use complex, noisy multimodal information to accurately infer meaningful things about the ongoing operation. Most importantly, these multimodal arrays may be expected to include data streams from human operators, which adds significant complexity. Importantly, the technologies and methods for achieving this kind of inference and prediction, particularly with humans in the mix, are rapidly evolving and thus persist in need for improvement.

Initially conceived in 2019, this project focused on the Army's needs that were discussed on the website of DEVCOM Army Research Laboratory's Center for Agent-Soldier Teaming (CAST)[†] as well as in the Strengthening Teamwork and Robust Operations in Novel Groups (STRONG) program (cf. DeCostanza et al. 2018). At the time, ARL researchers had been soliciting efforts on novel methods to reconceive the role of the human in the control and management of complex intelligent teams (e.g., Nothwang et al. 2016; Marathe et al. 2018) and were advocating strongly for individually adaptive technologies. The methods of interest were particularly focused on real-world applications of "neuro-technologies" for studying human decision and action (McDowell et al. 2013), such as brain activity (electroencephalography [EEG]) and bodily auto-regulatory mechanisms through cardiac (electrocardiography [ECG]) and skin electrical conductance. By the time she discovered ARL's research, Dr. Bhattacharya had already demonstrated success in using biomedically informed machine-learning methods (Bhattacharya et al. 2017) as well as analysis of bio-behavioral signals (brain activity, eye-gaze) to identify and predict onset of human driver intent and distraction (Bernadin et al. 2018; Bernadin and Bhattacharya 2019; Bhattacharya and Bernadin 2018, 2020;

* Contract no. W911NF2020205 (U.S. Army ACC-APG-RTP W91).

[†] Between FY2017 and FY2022, ARL hosted nine virtual research centers that were intended to facilitate dialogue between Army science and engineering programs and the broader relevant international communities; the CAST was among these, and during its time it generated significant opportunities for ARL (averaging, for instance, 22,000 monthly visitors in FY2021).

Bhattacharya 2019), and therefore she proposed a novel vision for multimodal fusion that could enable “human as a sensor” concepts and related success in ARL’s human-guided machine-learning research (Goecks et al. 2019).

1.1 Background

By 2019, society was experiencing a steady progression of advances in robotics, machine learning, big data, brain-computer interaction, and everyday wearable technologies. With each advance reshaping various sectors (and one another), their convergent potential was explored and demonstrated across a myriad of human-supporting application spaces. Innovators promised significant enhancements to productivity, safety, and efficiency, especially in otherwise dull, dirty, and dangerous jobs—civil police forces explored the integration of drones equipped with computer vision for crime prevention and intervention at a distance, whereas military ordnance disposal and urban hazardous material handlers investigated wearables and semi-autonomous robots to mitigate exposure potential in volatile situations. In civilian and military sectors, self-driving vehicle proponents were emphasizing that, one way or another, humans would soon trade the thrill of driving for the convenience and safety of being along for the ride. That is, in large swaths of future domestic transportation spaces, humans are expected to primarily be passengers while the cars drive themselves!

Certainly, there is no lack of information about human occupants of moving vehicles; however, traditional automobile studies have not focused so heavily on the passenger, where society currently needs the most information. Decades of research have yielded plenty of insight into car drivers, including their physiology, cognition, emotion, and social behavior; a summary of which is available in any number of published reviews and meta-analyses (Young et al. 2007; Caird et al. 2008; de Winter and Dodou 2010; Bengler et al. 2014). The *passenger experience*, however, has not been scrutinized and quantified nearly to the same degree. Therefore, the work of studying the passenger experience and behavior in real-world vehicles is essential for improving safety, enhancing human comfort and public trust, and promoting effective social integration of autonomous vehicles into existing civil and military transportation networks (Elbanhawi et al. 2015).

1.2 Army Research Context

Aiming to address the gap regarding passenger experience and social behavior in real-world vehicles, as well as to conduct an ecologically valid study of passenger trust and communication patterns under authentic risk, an ARL-led team devised

an experiment known colloquially as the “Real-World Driving (RWD)” study.* This was an outcome of the significant history of Army investment in simulated and closed-course driving research that systematically sought to emulate real-world conditions and task dynamics in applied, Soldier-focused research (Davis et al. 2008; McDowell et al. 2008; Kerick et al. 2009; Oie and Paul 2009; Metcalfe et al. 2010a, 2010b; Vettel et al. 2012; Gremillion et al. 2016). By 2017, the ARL researchers had suggested that human interbrain synchrony was potentially predictive of significant driving events (Garcia et al. 2017). However useful these results may have been, from an operational standpoint, they and others had yet to be verified in the real world.

Indeed, prior research had successfully used physiologically informed methods to predict driving risk and susceptibility to social influence (Falk et al. 2014; Cascio et al. 2015; Schmälzle et al. 2017). Therefore, across all aims, the team sought to understand passengers’ contextualized experience and social behavior as influenced by varied factors beyond in-drive interactions, including records of natural variability in sleep history, physical activity, and daily social interactions (Vettel et al. 2018). The RWD study instantiated a unique “moving laboratory” paradigm, which is illustrated on the next page. This experiment provided and captured rich passenger experience data that was enabled by an instrumented vehicle with an internal wireless network, as well as wearable and fixed-mount devices, including an experiment control and experimenter workstation (in the rear passenger seats) and separate mobile hubs for capture of physiological data from human-wearable, lightweight, wireless commercial sensors borne by both driver and passenger. The approach facilitated study of the social and psychological experiences of vehicle passengers during natural traffic, at highway speeds, and, in a limited way, provides insight regarding those who may soon become passengers in self-driving cars.

Figure 1 summarizes the study timeline. During intake (Figure 1A), participants provided informed consent and then completed surveys, including data about their general social behavior, personality, and risk preferences. Both participants completed pre-drive surveys about stress and trust as well, and then they carried out tasks specific to their role in the vehicle (Figure 1B). The passenger independently watched videos of staged news clips to be discussion topics later. On-road (Figure 1C), the passenger led discussion about the topics, and the driver controlled the pace of conversation for safety. Post-drive (Figure 1D), both participants completed further surveys and memory tasks about the discussion

*Metcalfe JS, Vettel J, Falk E. Assessment of intra- and interpersonal brain and behavioral dynamics of driver-passenger dyads during real-world driving [internal protocol document]. ARL Human Research Protections Program research protocol. Army Research Laboratory (US); 2017. Report No.: ARL-17-092.

topics; the driver was tested for recall of topics only, whereas the passenger concluded a recognition memory task as a paired control. Weeks later, the dyad returned for a second drive (Figure 1E), with their roles switched: they repeated the drive with the new passenger (former driver) and presented a fresh set of news clips.

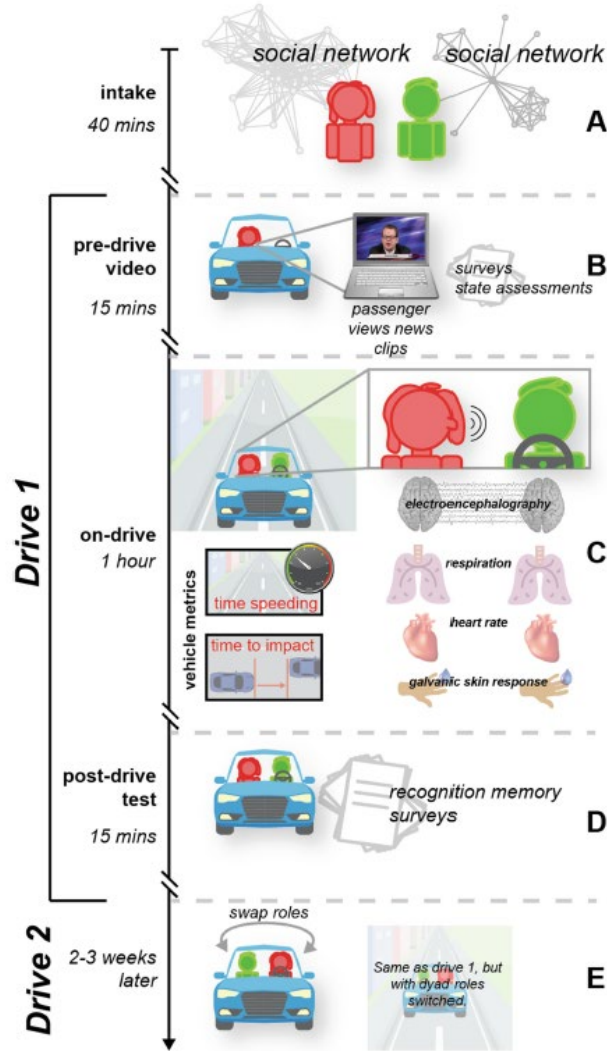


Figure 1. RWD study timeline and data elements.

By completion, the team had successfully conducted 43 of the 44 scheduled drives (one canceled for non-technical reasons), wherein each participant completed two open-highway sessions—once as passenger and again as driver. The resulting data set* contained over 130 h of driver-passenger data, including interpersonal memory and communications in open, cued, and recognition recall tasks, both on-road (passenger) and while stationary (driver). Each drive followed an approximate 50 mile and 1 h prescribed out-and-back path traversing both local and interstate

* A persistent and curated copy of the dataset is publicly available at https://dev.cancta.net/C3DS/db_login.php.

roadways in Harford County, Maryland. In addition to the communications, memory, and driving performance, physiological data were recorded from brain, heart, and skin electrical activity and later synchronized for potential assessment against more than 345,000 individual instances of over 308 distinctly logged event types. The present project focused on the passenger subset, particularly the physiological and behavioral responses of passengers to meaningful driving events.

Appendixes A–D are designed to highlight the people supporting this project, the University that has manifested the ecosystem supporting them, and of course, to review the documented outcomes of their efforts. Table 1 summarizes the content of each appendix.

Table 1. List of appendixes.

Appendix letter	Appendix title	Contents
A	About the Collaborators	Brief biographies about the organizational units and individuals involved in this research.
B	Student Participation	Non-identifiable student demographics relating to major fields of study, degree level, type of involvement, and related. Biographic highlights.
C	Research Dissemination	Data summary sheets highlighting 1) key elements of the seven publications resulting from this effort as well as 2) efforts that are expected to carry forward.
D	Technical Supplement	Passenger experience; multimodal fusion and human state estimation; and code catalog.

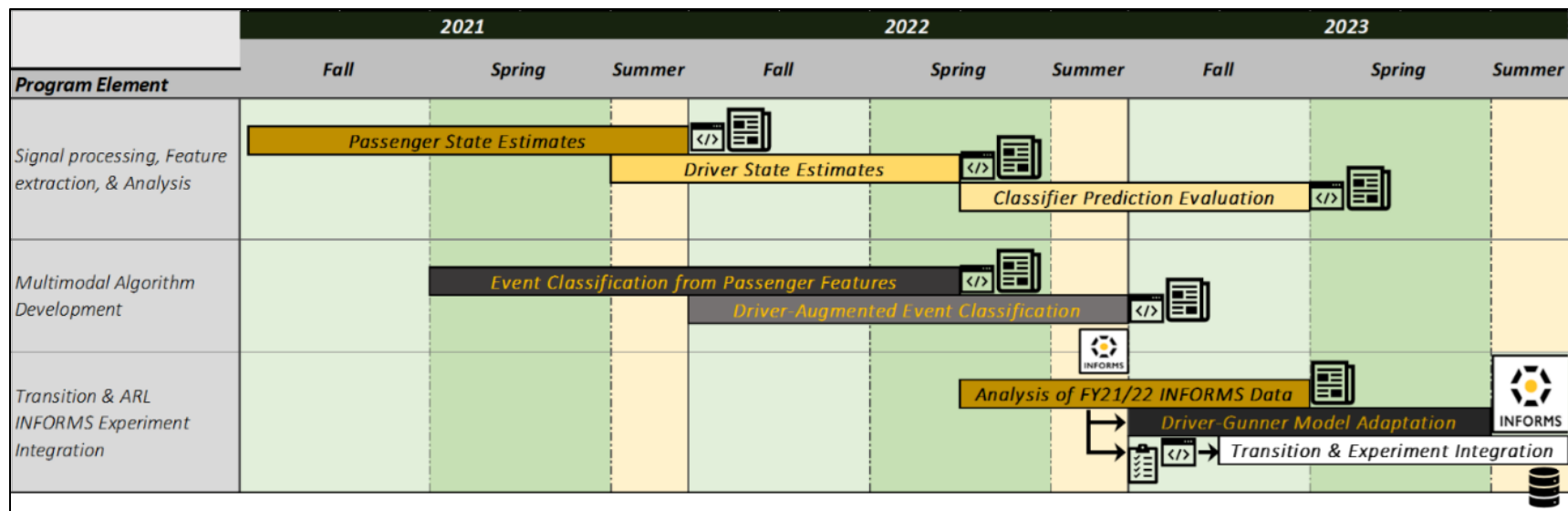
Appendix D provides a set of technical products that are direct and usable outcomes from the effort. These include the following: data regarding the real-world driving event context that are unpublished elsewhere, a technical description of the evolution and current state of the Sequential Event-based Multimodal Fusion framework, an outline of the vision for new empirical efforts to link these outcomes with developing methods under a separate ongoing effort also supported by cooperative agreement with U.S. Army Combat Capabilities Development Command Army Research Laboratory (“Overcoming Confusion in AI: Integrating Human-like ‘Gist’ into Complex Dynamically Evolving Human–AI Systems”; no. W911NF2320226; FY2024–FY2025, Principal Investigator [PI] Bhattacharya), and a catalog of the software developed in support of this project’s goals.

2. Project Overview

2.1 Initial Program Plan

Figure 2 is a depiction of the initial vision of the effort. The project was designed to develop useful methods for event detection from passenger behavior and physiology. In the second year, event detections were to be augmented with features from the driver. The ultimate vision was to translate driver-passenger models to driver-gunner situation modeling in the FY2023 Soldier experiment in the Information for Mixed Squads (INFORMS) demonstration and transition laboratory at Aberdeen Proving Ground (APG), Maryland.

7



Note: Background colors correspond with KSU academic terms. Icons: software (📄), reports (📋), INFORMS laboratory scenario specifications (🔍), and data (🗄️).

Figure 2. Initial program plan for FY2021, with FY2022 and FY2023 options.

2.2 Program Plan Development

The long-range objective of this research was to develop a robust method for integrating a multimodal set of sensors, including physiological, behavioral, and environmental, to enable reliable inference of human states and cognition during interactions with a vehicle-based driving autonomy. Initial efforts were to focus on human-only data recorded in ARL's RWD experiment. In the initial program year, ARL and KSU would collaboratively engage research efforts toward this goal through two specific aims*:

- 1) Establish and validate multimodal methods for characterizing states of human passengers in an automobile as a function of vehicle performance characteristics and environmental events.
- 2) Develop machine learning algorithms for systematic and informed classification of relevant passenger states with respect to the driver and the ongoing drive.

In optional successive years, the initial plan was to focus on extending successful results toward a multihuman, multiautonomy scenario by the third year. The first steps toward this would be taken in Program Year 2, through a driver-passenger analysis and classification that integrated cross-modality information streams, pending opportunity to work with INFORMS study data. Finally, successful outcomes from the first two years were to transition to ARL, with the ultimate vision to leverage multimodal driver-passenger algorithms toward an operational scenario examining driver-gunner roles in the FY2023 INFORMS Summer Study.

Given the initial budget, the effort was planned to be led by a full-time, post-doctoral researcher at KSU's Marietta campus, with periodic travel to APG. In response to difficulties encountered during the hiring process, we devised an alternative approach and relied on a tiered team of KSU students; some were university-funded assistants, others were graduate or undergraduate students on ARL contract funds (hourly or stipend), and a few undergraduates chose to volunteer their time and efforts in exchange for experience, education, and mentorship.

The alternative strategy was quite successful. In the initial year, the team collaboratively executed on the new program structure (see Section 3.1), catching up on data integration and then beginning to address each aim as (re)planned. First, the team established methods for characterizing human passenger states during

* Program plan for Cooperative Agreement between Kennesaw State University, Principal Investigator (PI) Dr. Sylvia Bhattacharya, and ARL, Dr. Jason S. Metcalfe, Cooperative Agreement Manager (CAM), concerning topic: multimodal inference of human state to track cognitive processes in risky environments.

roadway driving from multiple modalities. As will be detailed in successive sections, such methods included head and body pose and posture from vehicle-based cameras and seat-fixed force sensor arrays, as well as brain and cardiac electrical activity from surface measures. Second, machine-learning applications enabled classifying relevant passenger states in association with subjective judgments; for instance, examining potential associations between self-reported trust and brain-evoked potentials. Thereafter, second-year efforts aimed to expand the preliminary EEG results, to improve the robustness of the single-modality classifier performance and to begin the work of integration into multimodal features for novel fusion algorithms. The goal of the second year was to “integrate the multimodal features identified in the first year through novel fusion algorithms aimed at enabling inference across driver-passenger dyad in the spirit of ‘human as a sensor’ and ‘collective intelligence’ concepts.”*

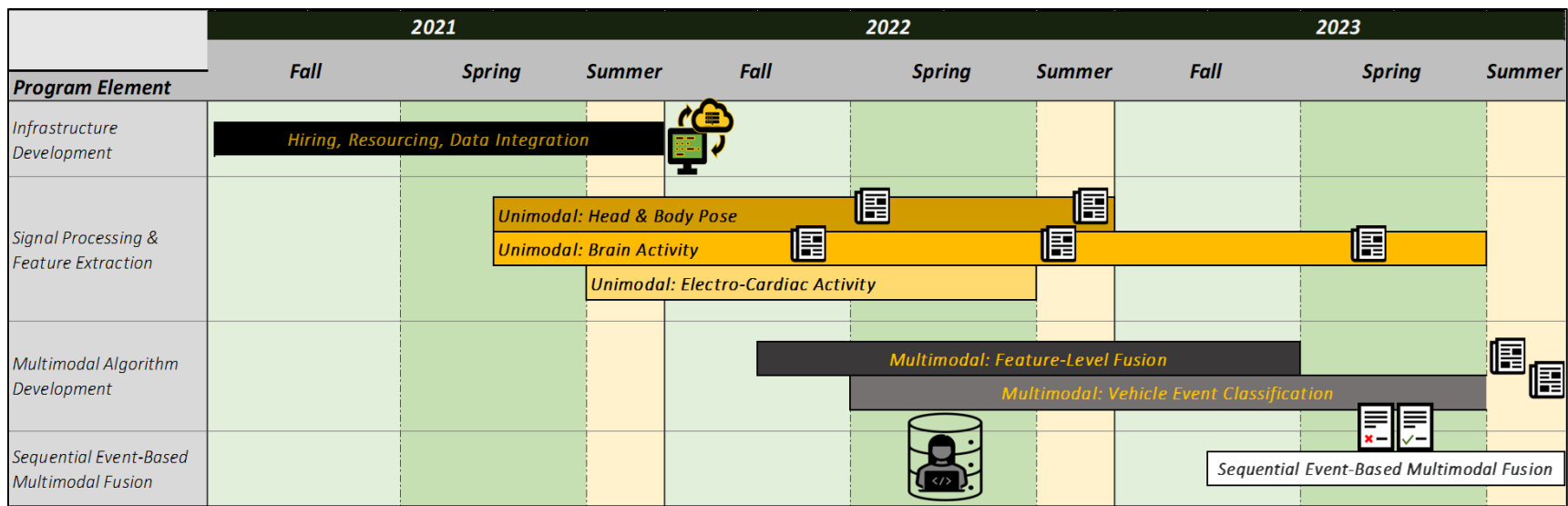
In the final year, owing to security-driven changes in external factors that dictate how Army-owned data of any type are to be shared, the data from the INFORMS laboratory were no longer readily available for analysis. Therefore, the originally intended transition from driver-passenger to driver-gunner models was supplanted by a more intense focus that fully fleshed out and built an integrative multimodal fusion framework.

3. Execution, Evolution, and Results

3.1 Executed Program Overview

Figure 3 depicts the project as it was finally executed. The project was modified from the initial vision for several reasons described previously. The adjusted vision was redesigned to build toward an improved framework and set of methods to deliver a multimodal data fusion solution, leveraging a broader set of skills across a newly formed team of undergraduate and graduate students.

* Program plan update for Cooperative Agreement among Kennesaw State University, PI Dr. Sylvia Bhattacharya, and ARL, CAM Dr. Jason S. Metcalfe, concerning topic: multimodal inference of human state to track cognitive processes in risky environments.



Note: Background colors correspond with KSU academic terms. Icons: KSU high-performance cluster integration of ARL data (), reports (), proposals (), and knowledge base ().

Figure 3. Executed program, FY2021–FY2023.

3.2 Overview of Outcomes

The 3-year collaborative effort culminated in a set of interdisciplinary outcomes that, although demonstrating advancement over present leading multimodal fusion methods with real-world data, also assemble into a solution applicable to the ARL RWD dataset as well as the others that share its constraints and statistical characteristics. The appendixes to this document incorporate a descriptive statistical table regarding the quantified passenger experience during data collection, unpublished elsewhere, as well as a catalog of software and algorithms applied during the probes of the RWD set. Importantly, each of the technical thrusts from this effort was PI directed, but *student championed*, producing ends that supported their education in mutually developed ways. As a collection, the products span empirical and conceptual thrusts. This section provides an overview, highlighting the most significant outcomes. Greater details are available in the appendixes.

3.2.1 Toward Data-Driven Insights into Passenger Experiences

After establishing personnel and computational infrastructure, as well as integrating a local working copy of the dataset on the high-performance computing cluster at KSU, the team rapidly learned to navigate and manage the complex dataset and built familiarity with its rich contents. Initial efforts focused unimodally on EEG and head pose. The first year thus laid the foundation for understanding the ways to explore the interplay between drivers' and passengers' brain and behavior during the on-road communication and memory task execution. Through a growing signal processing and machine classification toolset (see Appendix Table D-3), feature extraction methods were applied to multimodal data sources—namely, EEG, ECG, head position, and postural kinetics—and the team identified several tractable approaches to extracting useful indicators of passenger states during real-road driving scenarios. ECG data were available for application of known tools to retrieve estimates of heart rate and heart rate variability metrics using the HeartPy toolkit.* EEG data† were more immediately amenable for feature extraction, owing to the ARL Big Data for Human Sensing (ABDHS) EEG pipeline developed under the Cognition and Neuroergonomics Collaborative Technology Alliance (CaN CTA) (Tourneyan and Lee 2021). With the preprocessed data, several teams made

* <https://python-heart-rate-analysis-toolkit.readthedocs.io/en/latest/>

† All EEG data had been curated and pre-processed to Standard Level 2 as specified in the data description document available at <https://dev.cancta.net/C3DS/SANDR%20Data%20Description%20v2.2.2.pdf>

use of the established MNE toolkit* for Python analysts and transitioned all code to our common repository.

In one of our more revealing outcomes, the work of Belcher et al. (2022) provides a tantalizing example of the potential we see for these methods to enable a “human as a sensor” concept. We draw attention to humans’ unique capability to identify meaningful (vs. irrelevant) events during complex tasks. Of course, humans are unique in this ability because it is their priorities that define the relevance. Therefore, we assert that human bio-behavioral responses primarily reflect their evaluation of perceived events in relation to the meaning of those events to that person, within the situation. To further illustrate, consider the data presented in Figure 4. This figure provides time-resolved signal decompositions in the spectral domain, representing concentration of averaged global brain activity at different brain wave frequencies. The data are specifically for one driver–passenger dyad during three separate events: aggressive acceleration, braking, and lane change. These decompositions differ from D’Souza et al. (2021), who pioneered our team’s event-related approach using the Fourier Transform for estimating similar brain signal properties to compute traditional brain response measures. The Morlet wavelet that produced the results in Figure 4 is a complex exponential function multiplied by a Gaussian window (in this case, 500 ms wide); this yields a capability to recover spectral properties (signal strength at different frequencies) evolving over short spans of complex data. The technique was specifically well-adapted to the study of event-related activity during RWD, as is demonstrated by the data shown in Figure 4. Given this initial success, at highway speeds, it may offer an opportunity for building on previous Army brain–computer interface (BCI) technology successes in these domains, such as in the CaN CTA program.

* <https://github.com/mne-tools/mne-python> (also see Appendix Table D-5).

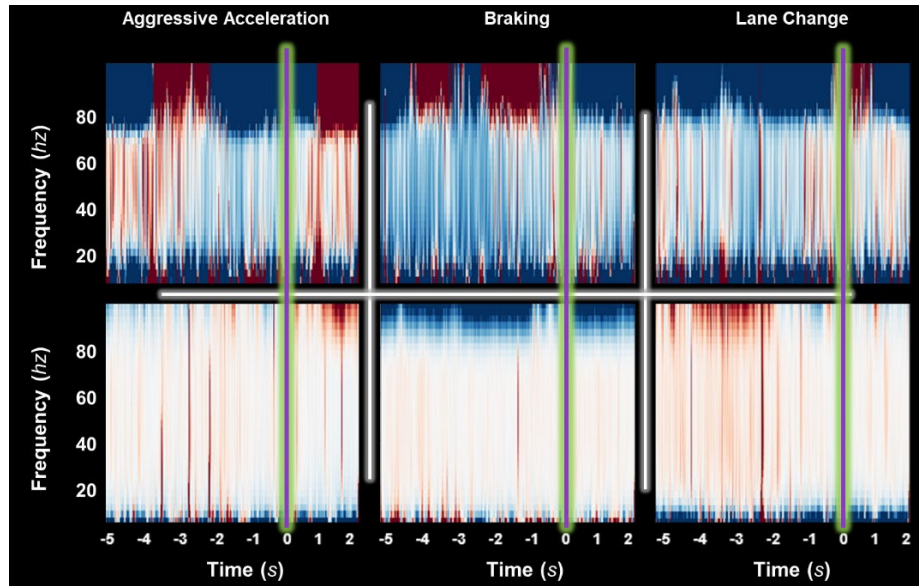


Figure 4. Average brain activity for a passenger-driver (upper-lower) dyad during three driving events.

The charts in Figure 4 are arranged to highlight the passenger. Thus, the top panels show the averaged global brain activity of the passenger during aggressive acceleration, braking, and lane changing (scale: red hues are higher values, and blue hues represent low values) (Figure 4). The bottom panels contain averaged global brain activity data for the driver, which is included to provide situated *face validity* for these measures (Figure 4). That is, a well-known EEG phenomenon is the *suppression* of global activity during periods of intensely focused concentration, usually accompanied by very focal activity in the lowest portions of the spectrum, as shown in Figure 4 for the driver, who is presumably focused on the task at hand throughout. However, please see our comment on the next page (end of first full paragraph) about what may be contrary to expectations, if not concerning, about this driver's EEG.

In the chart, the increments of frequency are labeled on the vertical axes, from very low to very high (0.1–100 Hz). The horizontal axis is event-relative time in seconds. A colormap scale was not provided in the original journal publication because the values were normalized in the averaging process: here, red hues are high (largely positive) values; blue hues are low (largely negative) values; and values in between lose saturation and become white as they pass through the normalized middle (“zero”). The horizontal lines indicate the event timing. The activity in the highest portion of the frequency spectrum of passenger is particularly relevant to our (and Army) research priorities. In the 50- to 80-Hz range, EEG is labeled “high gamma,” and high gamma has been related to various “meta”-processes believed to be involved in state regulation in cognitive and affective settings (Kang et al. 2014;

Yang et al. 2020); for instance, regulation is crucial for decision making under stress. Fluctuations in high gamma have long been considered useful for various empirical and theoretical reasons (Crone et al. 2011) and has recently been demonstrated to support classification of states into fundamental affective state elements in a multiclass environment (Gaertner et al. 2021).

The important part to us is the appearance of “bursts” in the highest frequencies (dark red patches) between 80 and 100+ Hz in the passenger charts. One may note that heightened activity periods each have a distinct time-based separation from the event that is on the order of seconds or less. For example, such periods occur approximately 3 s before and approximately 1 s after the onset of the aggressive acceleration in Figure 4. Likewise, there is a burst approximately 2 s before the braking event, accompanied by one earlier than that, at -4 s. The first one is *not* associated with any increase in activity for the driver at that same time, whereas the second one is. Finally, such a burst is nearly coincident with the lane change onset. Given the nature of the events, and considering the passenger’s general ability to anticipate them, we suggest the *possibility* exists that the presence of these bursts may have been meaningful and useful indicators of the passenger’s evaluation of the driver’s actions at the time as well as, perhaps, their own regulatory responses to observing internal or external driving conditions. To further emphasize the point, as it turns out, the most brain activity in the driver’s charts is largely confined to frequencies most associated with drowsiness and fatigue, accompanied by some indication of a possible suppression of high gamma before and through the braking event! Most of the time, drowsiness and suppressed regulation are not a good combination when a person should be calmly executing lane changes in U.S. East Coast traffic at highway speeds! If the passenger was aware of the driver’s state, or if the quality of their driving betrayed it, it may also be possible that the apparent high-frequency activity could be related to global state regulation.

Indeed, without knowing much more about the subjective aspects of the situation *and* about the relationship among the two people, we are unable to seriously consider such interpretations as described above. Likewise, unless the bursts can be shown to be systematically related to the driving events, across an appropriate population of people, with adequately powered statistics, we should *not consider any such narratives as standing on sure footing*. Instead, having offered the potential for such a phenomenon as being detectable, and having documented our method, we and others may explore its usefulness in future projects. Until then, we leave the more tantalizing conclusions to the domain of encouragement of our pursuits. We are most enthusiastic about the potential for applying these methods toward performance improvements in human-machine teams, especially through enhancing our understanding of the nuanced relationship between these potentially

observable process signatures with downstream decisions under risk and stress while processing a multisource information deluge.

Of course, an important way to immediately strengthen our ability to make use of results like Belcher et al.'s (2022) is to link the observed brain activity to a behavioral act of the person. If there was an indicator that the driver and passenger were both gazing at the same region of space during the lane change event, for example, there would be added evidence to strengthen inferences of that burst's relation to their experience of the drive at that moment—adding important potential semantic content that transforms complex analyses into our human understanding. Therefore, we next highlight a second success story: this one focused on how Kalidindi et al. (2022) used a rapidly developed and prototyped solution to address an issue that prohibited use of the RWD facial video camera data in raw form. The video data were originally intended to enable inference of gaze region and head displacement during the drives, ideally to enable identification of times of shared awareness by driver and passenger, but certainly these data also help with identifying and removing movement artifacts in physiological signals.

The challenge was caused most proximally by the custom cabin-affixed cameras; these were nonstandard pieces of equipment being applied beyond their original design context (web cameras), and their resolution and field of view were both limited and positioning was constrained by safety needs. The final views provided partial frames of the face for the driver and passenger, each from the opposite side of the vehicle to maximize field of view. Moreover, the view frame only encompassed a dynamic proportion of the participant's face that fluctuated between approximately 0.2 (or less) during an extreme turn away from the camera and 0.7 when looking partially at their co-participant but not directly at the camera. This was variable across passengers, depending on how much they naturally tended to move their head and shift their gaze during conversation in a moving car.

Figure 5 depicts the system that was designed to recover a set of minimally useful labeled coordinate ranges for machine classification of left, right, or center head positions. Without such a method, the videos, and the effort spent obtaining and managing them during the study, would have been lost. The solution relied on training a support vector machine (SVM; a machine classification technique) classifier to associate different partial frames of head and face views with moments when the passenger was identifiably looking in one of three directions: left, right, or center.

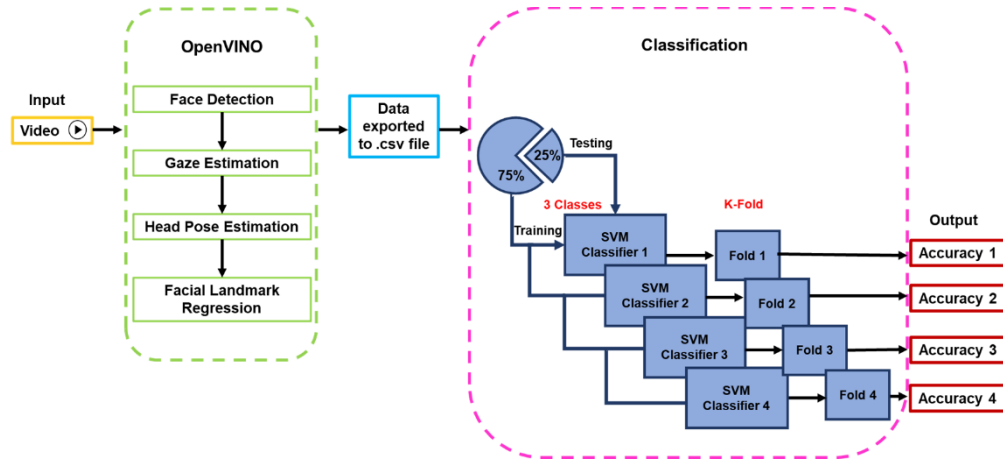


Figure 5. Custom system to extract head pose for each video frame.

The SVM assigned class membership to video frames based on features extracted with a custom-designed convolutional neural network ([CNN] a technique for visual image processing and decomposition), developed in an open-source commercial environment that was well suited to this task. After training, videos for 23 of the passengers were submitted to the process. These data allowed our team to make an initial determination that, depending on the participant, head positions were classified accurately, within a reasonable measure of confidence. Specifically, the head positions of 16 of the 23 were classified with approximately 90% accuracy or greater (Appendix Figure C-6). There were several participants for whom accuracies ranged between approximately 50% and 70%. This latter group likely represents an important hurdle to this specific method; that is, when participants did not move their heads very much naturally (perhaps), this machine-learning system failed to find any observations to classify as anything other than center. One mitigation for this is inclusion of other ways of estimating similar things with related modalities (such as the electro-oculogram, which may be obtained in several ways along with EEG).

In the end, the joint success of this head pose classifier and the Morlet-transformed EEG led our team to anticipate further success from the additional modalities (e.g., ECG and seat sensors). We thought that the added information would strengthen or otherwise augment estimates of the driving events. Ultimately, this was not the case, and our work led us to different and potentially more useful conclusions about multimodal fusion.

3.2.2 Future Multimodal System Concepts for Human–Machine Integrated Formations

Commencing from year two, our effort was divided into three parts: signal processing and feature extraction, theoretical systems integration concept

development, and multimodal algorithm development. The signal processing and feature extraction work continued, including a limited attempt to 1) improve the precision of EEG features as well as 2) add new data features related to whole-body motion based on seat- and chest-borne ECG sensors; however, there were no further significant advancements from this. An early event-related analysis for heart rate and heart rate variability saw some initial promise, but, for personnel reasons, that effort ended once basic rate and variability features were available for classification. The related functions and code fragments remain available in our GitHub* repository (see Appendix Section D.3). Similarly, a considerable effort was placed into classifying postural positions in a manner akin to the head pose (left, right, center). Although some early training runs were promising, results were not robust under cross-validation. As with ECG, the seat sensor exploratory software is available in our repository (see Appendix Table D-5); in this case, lessons were learned that inspired the student to focus on developing suggestions for system integration and design considerations related to this modality (Reddy Dendi et al. 2022).

In addition to Reddy Dendi's (2022) smart seat sensor review, our team produced a final and more integrative attempt at a full systems concept for EEG-based vehicle control (Biswas et al. 2023). The goal of both theoretical efforts was to assemble an overall systems concept necessary for the eventual engineering involved in implementing a fully multimodal classification system. Figure 6 was based on the work published by Biswas et al. (2023), but it expanded to incorporate all other results from this effort. Here, a complex Army circumstance is considered, demonstrating the potential for interfacing with uncrewed semi-autonomous unmanned ground vehicles (UGVs) and autonomous unmanned aerial vehicles (UAVs). Figure 6 visually encapsulates the BCI system that was the focus of Biswas et al. (2023). This image explicitly shows a human-machine integrated formation involving three ground vehicles and three drones, coordinated in a multiagent, multioperator BCI framework. The diagram is also annotated with citations to indicate which portions were informed by which of the different subteam efforts.

* <https://github.com/ARL-KSU>

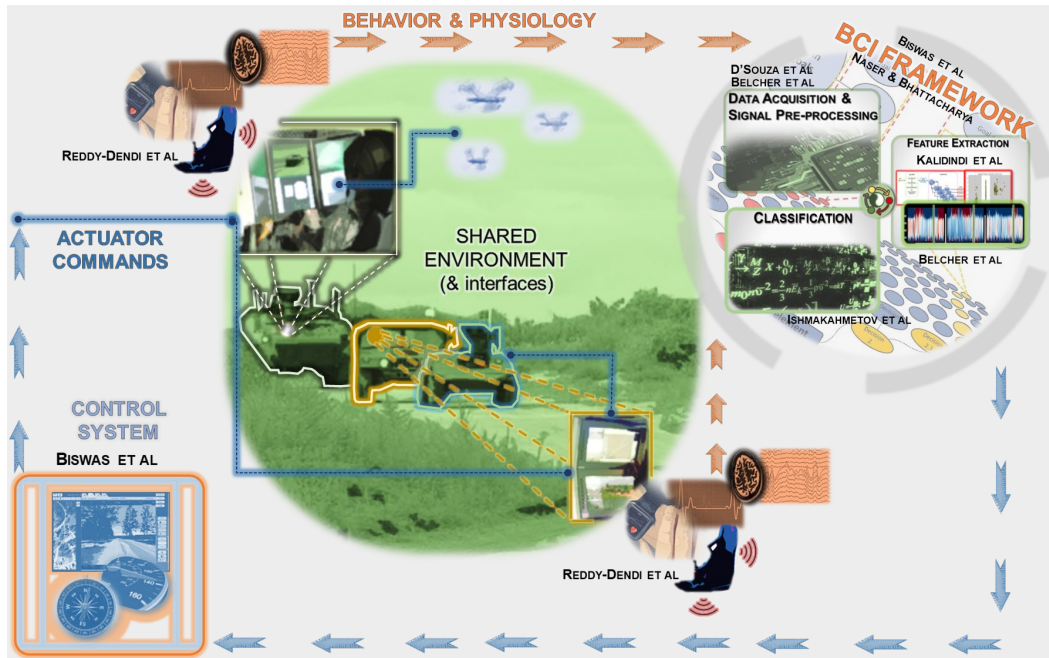


Figure 6. Conceptual brain-based control system (based on Biswas et al. 2023).

Figure 6 illustrates a military ground vehicle convoy through an environment depicted with single road through what appears to be dense grass and low desert vegetation. The image shows three ruggedized vehicles in a single-column formation, portrayed in a green-filtered “soda-straw” frame. Above the ground formation three, quadrotor-style UAV drones hover midair (Figure 6, blue-filtered for emphasis). The apparent road march is led by the UGV (Figure 6, light-blue highlight), which is immediately followed by an agile, all-terrain Humvee (Figure 6, light-orange highlight). At the rear is a 20-ton, 8-wheeled, light-armored Stryker (white highlight), partially hidden by the tall grass from which the convoy is emerging. Figure 6 contains many annotations, most obviously, the Humvee and Stryker are each associated with callout-style image insets that show the internal user interfaces for the operators of the UGV and UAV systems. Likewise, the interfaces are each depicted as a central part of the interconnected network of vehicles, sensors, and Soldiers by the blue actuator command lines that emanate from the control system represented with an array of interface screens and vehicle instruments, showing the closed-loop monitoring and information routing that is specified through an overall brain–computer interactive framework.

Figure 6 emphasizes integration of lightweight sensing, such as a wrist-borne ECG sensor, a light array of EEG electrodes, and an instrumented seat—much like that deployed in the RWD study. To enable the capture of data for this situation, robust and secure networking solutions in and out of each vehicle is required. Framed alternatively, each vehicle–human–asset combination may constitute a different

independent unit, essentially emulating dynamic, partly self-maintaining, endpoint wireless sensor networks. Dynamic networking, whether beamforming, mesh, or cloud, will therefore be critical to the BCI framework (central/distributed) that enables monitoring and use of resources; it also facilitates coordination across teams sharing these multimodal streams, particularly as those teams become increasingly distributed.

3.2.3 A Revised Multimodal Vision

In addition to facilitating an operational vision (as shown in Figure 6), Figure 7 highlights the dynamic interactions within the BCI framework, especially among—and within—the data conditioning, feature extraction, and classification processes. The internal workings of the BCI framework here differ from those shown in the original work (Biswas et al. 2023, p. 3). This vision incorporates dynamic interdependencies among modalities and processes in a complex, multiagent, and mixed-initiative context. In such situations, as detailed by Naser and Ishmakhametov in Appendix Section D.2, feature extraction and later classification algorithms must be fundamentally beholden to the reality of the event dynamics.

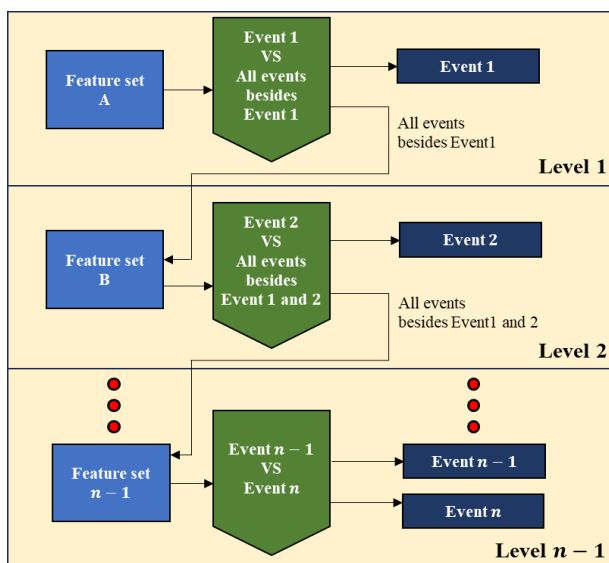


Figure 7. Sequential Event-based Multimodal Fusion ([SEM-F] developed in NIIL at KSU) yields ensemble outputs from hierarchical, multimodal feature-based decision layers.

To say that feature extraction and classification is beholden to event dynamics is to acknowledge important physical realities, such as the timing and nature of impinging physical energies, that are likely to differentially influence the data output from each sensor modality according to the design of that sensor. For example, EEG and ECG both use electrodes to detect changing micro-currents on the skin surface, but head motion in the RWD study was recorded through digital

video, and body adjustments were registered as changing body pressures on deformable force sensors woven into an aftermarket seat cover. Therefore, it most likely was not always the case that any given driving event was equally propagated into each sensor modality. Rather, it was more likely that certain events excited certain modalities or combinations thereof, and therefore our classifier system was designed to allow for variation in the modality or feature subsets that best identify specific events. Furthermore, this also implies that the fusion process may not have to continuously occur throughout all operations when deployed. Instead, it may best be deployed sparingly, as a means of enhancing data and energy efficiency for eventual real-world applications. Toward this end, our future efforts will pursue a better understanding of the joint questions: 1) “When or in which circumstances is the cost of multimodal fusion worth the value added?” and 2) “Once it is known that fusion is required, by which principles do we systematically choose particular data (e.g., modalities, epochs, and features) to fuse at that moment?”

4. Key Highlights, Values, Insights, and Outcomes

This Cooperative Agreement aligned ARL and KSU in establishing and nurturing a maximally collaborative research environment. As a result, varied contributions increased and extended beyond the originally anticipated science and engineering *products* as highlighted throughout. The final section of this reports casts these contributions in terms of value added. The contributions extend beyond the immediate advancements in technology and theory: they mark an important initial milestone for our ongoing efforts to advance multimodal methods for improved human–machine integration.

The incredible group of student scientists and engineers have grown along with this project, and they both provided a vision and initial roadmap toward advancing increasingly robust methods for exploring the complex interactions between humans and intelligent systems. The program’s efforts blended empirical research with conceptual review, design, and collaborative ideation, and the outcomes emphasize potential for increased interdisciplinary approaches to drive further progress in understanding and improving human–machine collaboration through the continued maturation of biophysical feature extraction methods. This work has expanded our multimodal classification capabilities and evolved our experimental approaches to yield advanced scientific paradigms and data that advance our understanding of how humans and machines cooperate in a multimodal world.

The finding and contributions are further documented in the annotated bibliographic datasheets in Appendix C. This compilation highlights the individual

studies within the program, and it contextualizes their collective impact, offering valuable insights for all who are involved in the lifecycle evolution of intelligent technologies that are meant to cooperate effectively, reliably, and safely alongside humans.

4.1 Ecosystem Growth, Diversification, and Talent Cultivation

Perhaps the largest values added were *growth* and *talent development* for 1) the students' future skills and 2) building a new generation of competencies needed to pursue ever-more advanced blends of human and intelligent machine! The laboratory-integrated student *experiences* across all levels of training enabled the successful *transition* of some of those same students into the science and technology (S&T) ecosystem at large, including in established industry and academic places.

By engaging 28 students from 7 disciplines, across 2 KSU colleges, majors from all levels (11 bachelors, 13 masters, and 4 doctoral) included the following: interdisciplinary, electrical, mechatronic and other engineering disciplines, data science, information technology, and computer science. Therefore, this project benefitted from the opportunity to engage with, learn from, and support emerging professionals with invaluable hands-on experience that was specific to their area of specialty. The project also provided an important complementary skill set that facilitated the expansive set of outcomes shown in this report.

Students also experienced personal challenge, success, and, most importantly, *opportunity* to transition theoretical, “textbook” learning (e.g., four biomedical modalities) into actionable applications (e.g., more than 3600 lines of analysis software). Through direct involvement in applying machine classification, physiological signal processing, and multimodal fusion, students were able to sharpen and contextualize their intellectual and professional *repertoire* as scientists and engineers while *advancing the S&T agenda* of the human–autonomy teaming for the Army. This experiential learning not only propelled their careers but also equipped them with the *technical expertise* to address complex challenges in their respective fields. For some, graduation was no boundary to continued generation of DoD-specific value.

Career outcomes for the participating students reflect the project's impact on talent development. Several students transitioned into new roles in industry (five companies) and academia (three program transitions at KSU and two acceptances to outside graduate programs). Notable successes include one student who won a place in a competitive internship program at Bose (Boston, MA) during the summer of 2024 and a former NIIL student who is currently employed with an ARL industry

partner and works on the APG installation. Finally, two students were accepted into graduate programs at competitive, top-tier institutions including Purdue (Doctor of Philosophy [PhD] program) and the University of Maryland (Master of Science program). This project’s commitment to diversity and inclusion further enriched the research environment, with a significant portion of participants coming from underrepresented groups (based on gender and ethnicity) in science, technology, engineering, and mathematics (STEM). This emphasis on inclusivity ensured that a broad range of perspectives and ideas contributed to the research, enhancing the overall quality and impact of the work.

The collaboration between ARL and KSU fostered a *sustainable* and *interdisciplinary* research ecosystem that extends beyond the immediate goals of the project. Moreover, the relationships between ARL scientists and the KSU students are direct and heavily focused on individual and group mentorship. At present, ARL is represented on two KSU dissertation committees and is in the process of formalizing an educational partnership with KSU. Indeed, this close collaborative environment not only facilitated the development of skilled researchers and young adult minds, but it also established a community that will *drive partnership and innovation* as well as offer challenges and contributions to changes in the practice and products of future science. The groundwork laid by this project has fomented a small *legacy*, affirming that KSU stands with its Army partners as leaders in human–autonomy teaming and advanced data processing for complex sociotechnical systems of the future.

4.2 Basic and Applied Research Contributions

Without question, KSU delivered on its promise to develop a new multimodal method capable of improving inferences about a risky task context from data derived from the human participants within that same context. This project made meaningful technical contributions by demonstrating the limits as well as offering pathways for addressing these limits with wavelet-based, time-frequency methods, and by offering a new integrated understanding of how to consider the development of multimodal data fusion at systems and algorithmic levels when involving physiological signal processing (see Figure 6). We argue that a significant advance lies within the proposal of the SEMF framework, and that the improvements we discovered may advance the process of fusing multiple modalities dynamically in complex informational environments (Appendix Section D.2). Indeed, although considerable developmental work lies in front of this team, we conclude this effort confident in our structured approach to integrating diverse data streams for human state inference.

The project also introduced several valuable insights into the application of existing techniques in a real-world-immersed experimental context. For example, the use of Morlet wavelets for EEG signal processing allowed the team to explore how cognitive neuroscience tools for studying evoked response potential (a stereotyped pattern of brain response to a discrete stimulus events) could be adapted for analyzing physiological signals in situ. In addition to the application of machine classification (i.e., SVM and CNN) to this complex, multimodal data provided a clearer understanding of the limitations and potentials in practice; Kalidindi et al.'s (2022) head pose estimator may prove to be both *highly deployable* and provide a significant *addition of informational value* for improved overall human and vehicle situational understanding. These efforts established the groundwork and provided demonstrations that can guide future research and development on complex multihuman (and multiagent) inference in dynamic multimodal informational environments.

The knowledge generated by the project was shared with the broader scientific community through seven publications (two journals and five conferences; one—the latest of which—inspired a conference workshop) that, to date, have accumulated a fair amount of attention, given the amount of time they have been available. This body of work that was produced through this project has been collectively cited 24 times in other works that were mainly published in respected, visible venues (median $h-5 = 27$, as indexed by Google Scholar) by authors who themselves displayed maturity in their research outputs (median citing, author $h = 12$, as indexed by Google Scholar). One hundred percent of these citations are from *external sources and research groups*, which adds visibility for the impact that the Army and ARL are generating on foundational approaches to advance technology integration.

Taken as a group, the delivery of an encouraging path forward for multimodal fusion *addresses several persistent challenges for signals and information processing*, as well as physiology-based machine learning applications as pertains to the leap over the “valley of death” into real-world application—not the least of which is how to adjust to dynamic changes in the quality (known or unknown) and trustworthiness of incoming information. SEMF and its related concepts are being carried forward into the exploration of the broader construct that we call *information arbitration*; that is, the adjudication and selection of the information upon which to base one's complex, fused inferences for risky decisions and actions.

4.3 Evolving Future Concepts to Deliver Transformational Overmatch

Ultimately, we believe that, through this project, our collaborative team established a solid foundation for future advancements in human–autonomy integration for dynamic real-world applications, aligning with the Army’s long-term goals of enhancing human–machine teaming capabilities. A key development from this work was the introduction of the SEMF framework, which offers a new approach to integrating diverse data streams for dynamic human state inference. Although SEMF is still in its early stages and requires further refinement, it has shown potential as a useful tool for making dramatic and near twice the improvements in the accuracy of preliminary classifications based on dynamic, multimodal information sets with any other state-of-the-art fusion approaches.

Finally, the cumulative effect of this 3-year project was to *lay the foundation for future growth* beyond this initial success and *expand research collaborations* toward communities based on future Army needs, as well as exploring broader applications of the developed methodologies into civilian health and educational sectors. The work on SEMF and related techniques has positioned KSU and its collaborators to *pursue further research and funding opportunities*, collaborative and independent, particularly in areas that require multimodal data integration. Although the focus within this initial project was primarily on human–autonomy interaction in roadway driving, the methods developed have potential applicability in other high-risk settings, such as military operations or healthcare. Here, we consider three specific future directions of high value for multimodal integration and fusion improvements.

- 1) *Data-Driven Human Factors in Autonomous Systems*: With an ability to gather important quantitative insights into the individualized and group dynamics of passenger trust in driving environments, we demonstrated the feasibility of using well-considered combinations of sensor modalities to identify data features to make predictions from observations of humans who are otherwise occupied in complex, high-consequence operations. In this way, this research offers a foundation for developing future systems that can better anticipate and respond to multi human states through the action of one or more intelligent, state-aware agents, if not actually make predictive environmental (or task) adjustments to effectively and proactively mitigate known consequences of impending group-level state changes, enhancing safety and confidence in autonomous technologies for improved future integration.

- 2) *Mitigating Classification Challenges via Multimodal Fusion*: The iterative development of classification algorithms, particularly the transition toward SEMF, reflects the NIIL and our collaborative commitment to overcome the inherent challenges of multimodal data integration for adaptive reasoning and judgment in very complicated operations. This focus on improving classification reliability directly contributes to risk mitigation in the deployment of autonomous and future AI-based systems, where accurate human state inference is critical.
- 3) *Strategic Extension of Research Collaboration*: Through this collaboration, a long-term vision has begun to emerge that appears to offer considerable potential for reshaping human-autonomy–AI (and other machine) groups in various ways. More importantly, our emerging vision emphasizes the importance of robust, adaptable, and context-aware systems that can operate seamlessly among, between, and interdependently with complex groups of humans in both close and distributed operations. Indeed, such a vision inherently points toward creation and broadening of scalability in research methods for complex cybernetic systems. Herein, the effort offered the beginnings of scalable methods, such as SEMF, for characterizing human passenger states based on multimodal data, including EEG, ECG, and head position. These methods have streamlined the process of human state inference in vehicle environments, laying a foundation for future research and potential applications in similar contexts.

5. Conclusion

In conclusion, this formal report and summary documents the formation, execution, and outcomes of the Cooperative Agreement between ARL and KSU. The 3-year project entitled Multimodal Inference of Human State to Track Cognitive Processes in Risky Environments* delivered the knowledge and technical products as collaboratively planned and offered novel directions upon which future collaborative expansion has begun. It is our hope that interested readers will find insight and inspiration when reading the remaining appendixes, which contain all technical and conceptual products from the program.

* Contract no. W911NF2020205 (U.S. Army ACC-APG-RTP W91).

6. References

- Belcher MA, Hwang I, Bhattacharya S, Hairston WD, Metcalfe JS. EEG-based prediction of driving events from passenger cognitive state using Morlet wavelet and evoked responses. *Trans Eng.* 2022;8:100107. <https://doi.org/10.1016/j.treng.2022.100107>
- Bengler K et al. Three decades of driver assistance systems: review and future perspectives. *IEEE Int Transp Syst Mag.* 2014;6(4):6–22.
- Bernadin S, Bhattacharya S. Effect of conversation on driving: towards increased mental load and virtual lateral vehicular variation. *Technol Interf Int J.* 2019.
- Bernadin S, Ehtemami A, Bhattacharya S, Scott R. Assessing driver performance of lane-merge tasks using video capture method. *Proceedings of the 2018 SoutheastCon*; 2018; St. Petersburg, FL. p. 1–6. <https://doi.org/10.1109/secon.2018.8479290>
- Bhattacharya S. A novel approach for AI-based driver behavior analysis model using visual and cognitive data [dissertation]. Florida State University; 2019.
- Bhattacharya S, Bernadin S. Eye-glance frequency as a function of driver's intent to change lanes. *Proceedings of the 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*; 2018; Chicago, IL. p. 1–5. <https://doi.org/10.1109/VTCFall.2018.8691029>
- Bhattacharya S, Bernadin S. Correlation studies on the cognitive responses of distracted drivers during conversational tasks. *Proceedings of the 2020 SoutheastCon*; 2020; Raleigh, NC. p. 1–5. <https://doi.org/10.1109/southeastcon44009.2020.9249678>
- Bhattacharya S, Bhimraj K, Haddad RJ, Ahad M. Optimization of EEG-based imaginary motion classification using majority-voting. *Proceedings of SoutheastCon*; 2017; Concord, NC. p. 1–5. <https://doi.org/10.1109/SECON.2017.7925328>
- Biswas S, Hairston WD, Metcalfe JS, Bhattacharya S. EEG based BCI for autonomous control: a review. *SoutheastCon*; 2023; Orlando, FL. p. 827–832. <https://doi.org/10.1109/SoutheastCon51012.2023.10115194>
- Caird JK, Willness CR, Steel P, Scialfa C. A meta-analysis of the effects of cell phones on driver performance. *Accid Anal Prev.* 2008;40(4):1282–1293.

- Cascio CN et al. Buffering social influence: neural correlates of response inhibition predict driving safety in the presence of a peer. *J Cogn Neurosci*. 2015;27(1):83–95.
- Crone NE, Korzeniewska A, Franaszczuk PJ. Cortical gamma responses: searching high and low. *Int J Psychophysiol*. 2011;79(1):9–15.
- Davis J, Animashaun A, Schoenherr E, McDowell K. (2008). Evaluation of semi-autonomous convoy driving. *J Field Robot*. 2008;25(11–12):880–897.
- DeCostanza AH et al. Enhancing human–agent teaming with individualized, adaptive technologies: a discussion of critical scientific questions. *IEEE Brain*; 2018. <https://brain.ieee.org/brain-storm/enhancing-human-agent-teaming>
- de Winter JC, Dodou D. The driver behaviour questionnaire as a predictor of accidents: a meta-analysis. *J Safety Res*. 2010;41(6):463–470.
- D’Souza KM, Dang T, Metcalfe JS, Bhattacharya S. Brain-based indicators of passenger trust during open-road driving. 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall); 2021; Norman, OK. p. 1–6. <https://doi.org/10.1109/VTC2021-FALL52928.2021.9625414>
- Elbanhawi M, Simic M, Jazar R. In the passenger seat: investigating ride comfort measures in autonomous cars. *IEEE Int Transp Syst Mag*. 2015;7(3):4–17.
- Falk EB et al. Neural responses to exclusion predict susceptibility to social influence. *J Adolesc Health*. 2014;54(5):S22–S31.
- Gaertner M, Sauter D, Baumgartl H, Rieg T, Buettner R. Multi-class emotion recognition within the valence-arousal-dominance space using EEG. Paper presented at: 27th Americas Conference on Information Systems; 2021 Aug 9–13; Virtual Conference.
- Garcia JO et al. Estimating direction in brain-behavior interactions: proactive and reactive brain states in driving. *NeuroImage*. 2017;150:239–249.
- Goecks VG, Gremillion GM, Lawhern VJ, Valasek J, Waytowich NR. Efficiently combining human demonstrations and interventions for safe training of autonomous systems in real-time. *arXiv*: 2019 July. *arXiv*:1810.11545v2. <https://doi.org/10.48550/arXiv.1810.11545>
- Gremillion GM et al. Analysis of trust in autonomy for convoy operations. *Proceedings of the Micro-and Nanotechnology Sensors, Systems, and Applications VIII*, Vol. 9836; 2016; Baltimore, MD. SPIE Defense + Security. <https://doi.org/10.1117/12.2224009>

- Kalidindi RS, Koonjah Y, Belcher M, Metcalfe JS, Bhattacharya S. Sub-optimal head position classification using OpenVINO and SVM. SoutheastCon 2022; 2022. <https://doi.org/10.1109/SoutheastCon48659.2022.9763965>
- Kang JH, Jeong JW, Kim HT, Kim SH, Kim SP. Representation of cognitive reappraisal goals in frontal gamma oscillations. PLOS One. 2014;9(11):e113375. <https://doi.org/10.1371/journal.pone.0113375>
- Kerick SE, Oie KS, McDowell K. Assessment of EEG signal quality in motion environments. Army Research Laboratory (US); 2009. Report No.: ARL-TR-4866.
- Marathe AR et al. The privileged sensing framework: a principled approach to improved human–autonomy integration. Theo Iss Ergon Sci. 2018;19(3):283–320.
- McDowell K, Nunez P, Hutchins S, Metcalfe JS. Secure mobility and the autonomous driver. IEEE Transactions on Robotics. 2008;24(3):688–697. doi:10.1109/TRO.2008.924261
- McDowell K et al. Real-world neuroimaging technologies. IEEE Access. 2013;1:131–149.
- Metcalfe JS, Alban J, Cosenzo K, Johnson T, Capstick E. Field testing of teleoperation versus shared and traded control for military assets: an evaluation involving real-time embedded simulation and Soldier assessment. Proceedings of SPIE. SPIE Defense, Security and Sensing; 2010a; Orlando, FL. 7692:769206. doi:10.1117/12.850597
- Metcalfe et al. Experimentation and evaluation of threat detection and local area awareness using advanced computational technologies in a simulated military environment. Proceedings of SPIE. SPIE Defense, Security and Sensing; 2010b; Orlando, FL. 7692:769209. doi:10.1117/12.850516
- Naser MYM, Bhattacharya S. Empowering human–AI teams via intentional behavioral synchrony. Front Neuroergon. 2023;4. <https://doi.org/10.3389/fnrgo.2023.1181827>
- Nothwang WD et al. Multi-sensor fusion architecture for human–autonomy teaming. Proceedings of 2016 Resilience Week; 2016 Aug. p. 166–171. IEEE.
- Oie KS, Paul V. The utility of ride motion simulation in a neuroergonomic approach to systems design. Proceedings of the 2009 Ground Vehicle Systems Engineering and Technology Symposium (GVSETS); 2009 Aug 18–20; Detroit, MI.

- Reddy Dendi PK, Gurjala Y, Bhattacharya S, Metcalfe JS. A review study of smart vehicle seat sensor for real-time postural analysis. *Proceedings of International Conference on Computing and Communication Networks: ICCCN 2021*; 2022 July. p. 229–241. Springer Nature Singapore.
- Schmälzle R et al. Brain connectivity dynamics during social interaction reflect social network structure. *PNAS*. 2017;114(20):5153–5158.
- Touryan J, Lee T. Cognition and neuroergonomics collaborative technology alliance final report. DEVCOM Army Research Laboratory (US); 2021 July. Report No.: ARL-SR-0450. <https://apps.dtic.mil/sti/pdfs/AD1144803.pdf>
- Vettel J et al. Mission-based scenario research: experimental design and analysis. *Proceedings of the Ground Vehicle Systems Engineering and Technology Symposium*; 2012.
- Vettel JM et al. Translating driving research from simulation to interstate driving with realistic traffic and passenger interactions. In: Cassenti D, editor. *Advances in Human Factors in Simulation and Modeling*. AHFE 2018. *Advances in Intelligent Systems and Computing*; vol. 780. https://doi.org/10.1007/978-3-319-94223-0_12
- Yang K et al. High gamma band EEG closely related to emotion: evidence from functional network. *Front Hum Neurosci*. 2020;14:89. <https://doi.org/10.3389/fnhum.2020.00089>
- Young K, Regan M, Hammer M. Driver distraction: a review of the literature. *Distracted Driving*. 2007:379–405.

Appendix A. About the Collaborators

Kennesaw State University (KSU)* is a comprehensive American university, located on two suburban campuses in Kennesaw and Marietta Georgia, in the Atlanta metropolitan region (Figure A-1). With approximately 45,000 students enrolled across 180 different undergraduate, master's, doctoral degree, and certificate programs, Kennesaw is the third largest university in the state of Georgia. Value-centered KSU places primary importance on being student first, research-driven (Carnegie-designated R2), and authentically diverse. As a full-service university, KSU supports a thriving residential population exceeding 5200. The highly engaged student body collectively leads more than 300 organizations across the university. Kennesaw's location in the Atlanta region provides access to a talented, diverse student body of domestic and international origins, and KSU places the highest priority on serving and connecting with this fantastic group of young minds from a wealth of perspectives. In 2021, KSU was ranked 14th nationwide for bachelor's degrees conferred to African American students, 75th for bachelor's degrees conferred to all underserved groups in all disciplines, and top 100 for master's degrees conferred to African American students. Indeed, KSU is more than a campus, it is an intellectually and culturally diverse Atlanta microcosm that is nurturing the next generation of innovators and thought leaders.

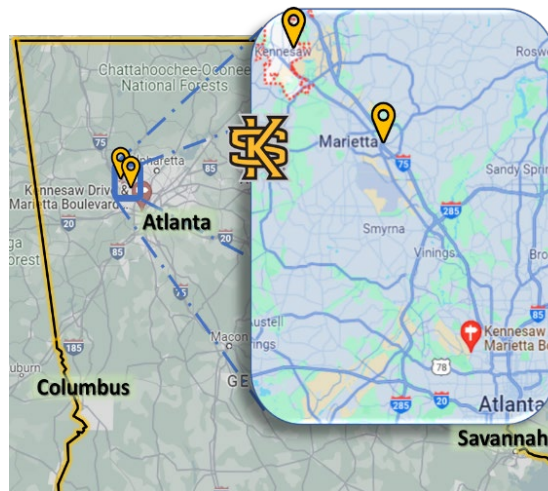


Figure A-1. Location of KSU campuses in the greater Atlanta, Georgia area.

The Southern Polytechnic College of Engineering & Engineering Technology, with its origins tracing back to March 1948 as *The Technical Institute* in response to the post-World War II (WWII) demand for technicians, has evolved significantly through the years. The institution was initially founded with 116 students, predominantly WWII veterans, and it was quickly recognized as a college-level school by the U.S. Department of Education after being renamed the Southern

* Read more at: <https://www.kennesaw.edu/about/>

Technical Institute. It underwent several transformations, including relocation to Marietta, Georgia in 1961, accreditation as a 4-year college in 1970, and achieving independence from Georgia Institute of Technology in 1980. After becoming Southern Polytechnic State University (SPSU) in 1996, it was consolidated into KSU in 2015 as the Southern Polytechnic College of Engineering & Engineering Technology, in memory of the legacy of SPSU. This progression reflects the institution's deep-rooted commitment to innovative education advancement and building on significant history with the U.S. government and the federal educational landscape.

The Neuro-Interaction Innovation Laboratory students say “Never stop neurovating,” a motto that is borne of the ethos that Dr. Sylvia Bhattacharya has cultivated since 2019, when she joined the faculty of KSU and established her laboratory. Although Dr. Bhattacharya has already achieved a significant number of accomplishments during that same time, she aptly notes that the true power of the laboratory comes from the strong culture of student leadership, co-mentorship, and team camaraderie (Figure A-2). With students at all levels, from undergraduate to doctoral, there is a highly collaborative air that accompanies their presence. Indeed, their welcoming collaborative energy extends to other laboratories and departments within the KSU ecosystem. Most commonly, the laboratory involves students from many domains of engineering, including electrical, mechatronic, and interdisciplinary specializations, as well as frequently integrates members from data science and analytics and computer science departments of KSU. Dr. Bhattacharya's laboratory explicitly focuses on supporting students from gender, ethnic, religious, and other cultural populations that are typically underrepresented in science, technology, engineering, and mathematics (STEM), with more than 85% fitting such demographics in any given academic year.



Figure A-2. Members of the Neuro-Interaction Innovation Laboratory in 2023.

Keyword descriptions of the research in this laboratory include autonomous and intelligent transportation systems, neural signal processing, human–autonomy teaming, brain–computer interaction, and multimodal machine learning. The

concept of “neurovation,” within this team, represents more than applying thought toward innovation, and rather acknowledges and invites the challenge of finding new ways of addressing the human brain and behavior through biomedical engineering. Under the astute direction and thoughtful mentorship of Dr. Bhattacharya, the students of this laboratory are currently focused on the following: 1) understanding interactions between humans and engineered systems (e.g., autonomous wheelchair and simulated autonomous driving) in both physical and virtual reality through application of state-of-the-art methods in biomedical engineering, computer, cognitive, and data science; and 2) evolving methods for processing and drawing complex task and environmental inferences from human brain and body signals for supporting more natural human–AI collaboration and partnership.

Sylvia Bhattacharya^{*} is a distinguished, multidisciplinary engineer at KSU and beyond, and she specifically stands out among her contemporaries in biomedical engineering (Figure A-3). Her goal is to pioneer advancements in multimodal data fusion and machine learning to improve Human Autonomy Teamwork and further advances in biomedical applications. Her expertise is in advanced machine learning methods toward classifying functional states and dynamics to enhance intelligent driving assistance algorithms. Dr. Bhattacharya’s career is marked by contributions to education, research, and through a 2-year fellowship role with Google AI, and contributions to real-world innovation at the cutting edge of industry.



Figure A-3. Sylvia Bhattacharya, PhD.

After obtaining her undergraduate biomedical engineering degree from West Bengal University of Technology in 2013, she moved to the United States for her Master of Science in Applied Engineering Degree at Georgia Southern University. In 2016, she commenced her doctoral studies in electrical engineering at Florida State University, contributing to research on driving distraction under Dr. Shonda L. Bernadin at Florida A&M University. After completing her PhD, she took a

^{*} Read more at: <https://engineering.kennesaw.edu/sylvia-bhattacharya.php>

position at KSU, where she has been instrumental in contributing to KSU's growth in research and innovation ever since. Currently, Dr. Bhattacharya is an Associate Professor of Engineering Technologies and serves the KSU community as the Faculty Champion Applied Innovation within the Office of Research. Her commitment and dedication are highly regarded by her students, peers, and leadership.

Appendix B. Student Participation

This appendix provides non-identifiable data about students who participated throughout the project.* Table B-1 shows each student by major area of study and degree level. The remaining table elements outline involvement and status. The following pages provide additional facts about the students as a group and on and an individual level.

Table B-1. Student data by major area of study and degree level.

		Degree Level			Involvement Status					Funding			
		Bachelors	Masters	Doctoral	2021	2022	2023	Graduated	Employed	GRA (ARL)	GRA (KSU)	Hourly	Volunteer
Major area of study													
01	Software Engineering		•			•		•	•	•			
02	Software Engineering		•				•		•	•			
03	Electrical Engineering	•						•	•	•		•	
04	Electrical Engineering		•							•			
05	Data Science & Analytics	•			•				•	•			
06	Electrical Engineering	•						•	•			•	
07	Data Science & Analytics			•						•			
08	Computer Science		•				•	•		•			
09	Mechatronics Engineering		•				•					•	
10	Electrical Engineering	•						•				•	
11	Software Engineering		•			•				•			
12	Data Science & Analytics		•		•					•			
13	Information Technology		•		•			•		•			
14	Interdisciplinary Engineering			•							•		
15	Electrical Engineering	•						•				•	
16	Data Science & Analytics			•						•			
17	Electrical Engineering	•				•						•	
18	Computer Science		•					•	•	•			
19	Computer Science		•		•			•		•			
20	Information Technology		•				•	•		•			
21	Data Science & Analytics			•							•		
22	Electrical Engineering	•					•					•	
23	Computer Science	•					•					•	
24	Electrical Engineering	•			•			•				•	
25	Electrical Engineering	•					•						•
26	Electrical Engineering	•			•			-	-				•
27	Computer Science		•		•			•	•	•			
28	Data Science & Analytics		•		•			•	•		•		

Notes: Connected horizontal lines indicate continuity; gray font indicates current Neuro-Interaction Innovation Laboratory (NIIL; led by Dr. Bhattacharya) students.

The NIIL is a cross-university laboratory that maintains a diverse talent base. Figure B-1 represents laboratory members from two Kennesaw State University (KSU) colleges, and the breakout chart at the right provides the number of students by major area of study and degree level. In addition to breadth of technical skills and academic level, the NIIL students also represent a set of intellectual perspectives built from a vast set of lived experiences. NIIL students represent a

* Note: All student spotlights were invited during the generation of this report. All contributions were voluntary and frequently returned with words of gratitude. Minor edits were made by the lead author.

global population, coming from South Korea, India, the Kyrgyz Republic, Jordan, Ethiopia, and the United Kingdom, as well as an excellent group of American students, including from North and South Carolina, Maryland, and, of course, Georgia (Figure B-2).

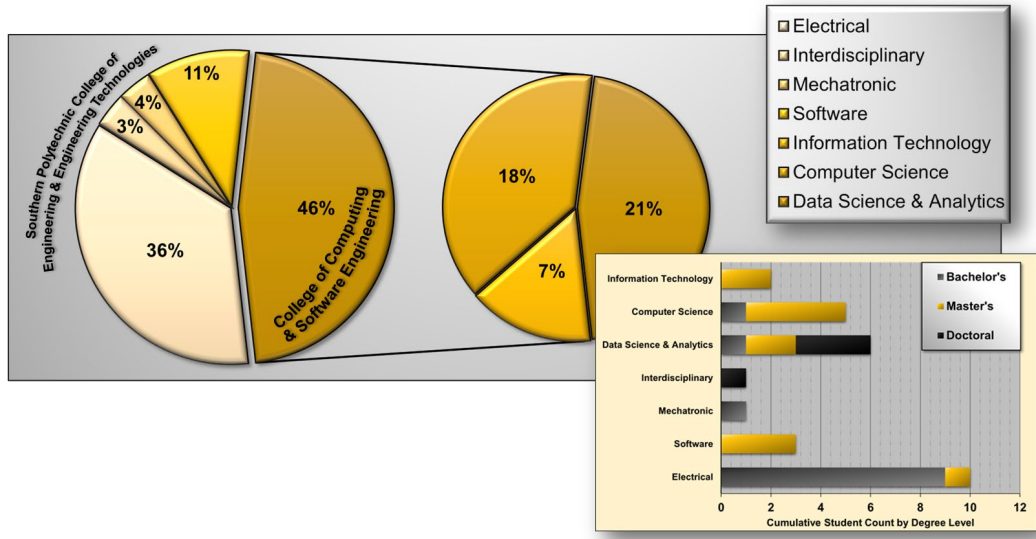
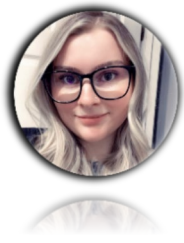


Figure B-1. Distribution of NIIL student major areas of study.



Figure B-2. Map of places of origin of NIIL student members.

B.1 Alumnus Spotlight Morgan Belcher



Morgan Belcher earned her bachelor's degree in electrical engineering from KSU in 2022 and was one of the foundational members of our collaborative team. By working in Dr. Bhattacharya's laboratory, she gained invaluable research experience that prepared her for a rigorous graduate program. She independently developed and managed the U.S. Army Combat Capabilities Development Command Army Research Laboratory-KSU code repository on GitHub until her graduation. She completed a summer student internship at Aberdeen Proving Ground, Maryland in 2022 under Dr. Hairston and worked on analyzing in-ear electroencephalography (EEG) data. She has matriculated to University of Maryland (UMD), where she completed her Master of Science (M.Sc.) in electrical engineering with a concentration in signal processing during the summer of 2024. Her research at UMD centered on auditory stream segregation and exploring methods to measure it without relying on linguistic content. She worked extensively on the collection of simultaneous magnetoencephalography (MEG), EEG, and pupillometry data, and she is leading the analysis of the MEG and EEG data for this project:

My research at KSU meant so much to me and really ignited my passion for research, and I carry that with me to this day.

B.2 Alumnus Spotlight Thomas Dang



Thomas Dang was part of the inaugural team that set up the high-performance computing cluster server and database for the real-world driving (RWD) dataset. He partnered with Karen [D'Souza] to publish the laboratory's first conference paper, which gave him a deep understanding of physiological signals like EEG and the complexities of preprocessing them. Before joining the [NIIL], Thomas had several internships in power distribution, manufacturing, and quality assurance, but he was not strongly drawn to continued growth in those domains; rather, he claims that his time in the laboratory attuned him to the constant engagement that research provokes in exploring, if not demanding, new methods and findings that can change the way we think. Thomas is now at DCS Corporation and provides engineering support directly to the neuroscientists at DEVCOM Army Research Laboratory at the historic Aberdeen Proving Ground, Maryland. Thomas professionally engages every stage of research, from experimental design and hardware and software validation to data collection, analysis, and visualization. He has presented his research on motion artifacts at conferences hosted by IEEE, the Society for Neuroscience, and MoBI

(mobile brain-body imaging), and he is currently preparing a manuscript for journal submission on the matured results. By working in NIIL, Thomas was able to lay a solid foundation for the next phase of his career, and he is now pursuing an M.Sc. in Machine Learning as needed for future goals of the military. Thomas credits the experience he gained in the laboratory as key to his successful acceptance into his present graduate program:

Working with Dr. Bhattacharya was crucial in my journey to becoming a biomedical engineer. It's incredibly rewarding to know that our discoveries have a tangible impact on the field.

B.3 Alumnus Spotlight Karen D'Souza



With prior engineering industry experience, Karen D'Souza gave crucial leadership in developing the initial EEG methods with Thomas Dang. Their efforts produced the first EEG paper from this effort that was accepted to any conference. In reflection, Karen expresses pride in experimenting with EEG that matched human behavior from a neuroscience lens, because this provided novelty, challenge, and excitement to her research. Finding inspiration from the regular brainstorming and mentoring moments, Karen was adept at converting identified gaps and blind spots, whether analytic or in presentation, into technical strengths. After these experiences, Karen fully invested in an applied machine learning research career; enthusiastic to discover more real-world settings that encourage neuro-inspired innovation. With her successful completion of an M.Sc. in Information Systems at KSU, Karen is presently a Computer Science doctoral candidate at Purdue University. She has also engaged several research internships in industry at Dell, IBM Thomas J. Watson research center, and Pacific Northwest National Laboratory centered around generative AI. Karen's research on multimodal modeling and human feedback may be found in venues such as Tapia, IEEE, AAI, and NeurIPS conferences:

The support, freedom, and guidance during this project were invaluable; I observed and experienced how machine learning research provides a crucial collaborative space to enrich cross-disciplinary biomedical research. My role on this project supported my growth profoundly, as I gained clarity on pursuing an applied machine learning research career.

B.4 Graduate Student Spotlight Mohammad Y. M. Naser



Mohammad Naser is a third-year PhD student at KSU, pursuing a degree in Interdisciplinary Engineering with a focus on Biomedical Systems. His background primarily includes Electrical Engineering (signal processing) and Numerical Modeling. This project has provided Mohammad with real-life project experience from the start of his program. This support has not only enhanced his skill set in physiological signal processing, but it also opened new avenues for innovation. These opportunities have been integral to his dissertation, which centers on developing a brain-based wheelchair control system using EEG and functional near-infrared spectroscopy (fNIRS) signals. This project has also enabled Mohammad to host a workshop at the IEEE Southeast Conference on the rigor of machine learning research in physiological signal processing, co-hosted by Namaz Ishmakhametov and Sylvia Bhattacharya. In addition, this work has inspired Mohammad to file a full utility patent, as well as secure an internship at Bose Corporation, a leading American company in audio signal processing. Alongside Namaz Ishmakhametov, Mohammad is a primary architect of a new multimodal fusion framework (Sequential Event-based Multimodal Fusion [SEMF; developed in NIIL at KSU], described in Appendix Section D.2):

As the only participant [in the 2024 Bose] internship who is from a non-R1 school, with other interns hailing from the top U.S. schools, I see how my niche experience, gained in part through this project, has been essential in developing and distinguishing my unique potential.

B.5 Graduate Student Spotlight Namazbai Ishmakhametov



Namaz Ishmakhametov is a third year PhD student in Data Science and Analytics, and he is currently leading the data-driven multimodal fusion modeling and integration methods development. Before starting his PhD at KSU, Namaz obtained a Bachelors-equivalent (Kyrgyz “Specialist”) degree in Mathematical Methods in Economics, as well as Masters in Analytics. Namaz worked as Data Scientist at Deloitte Consulting LLP in New York City where he served various Fortune 100 clients in the financial, retail, and healthcare industries. Namaz also spent 5 years as an econometrician the Central Bank of the Kyrgyz Republic, his country of origin, where he built predictive and explanatory macroeconomic models to advise on interest rate policy decisions in the context of a highly volatile national economy. With a deep interest in behaviors of intelligent collectives, such as markets, especially in periods of crisis that lead to

panic, and successively, poor decision making, Namaz was motivated to undertake quantitative study at the intersection of human emotion, judgment, and decision making. Ultimately, these interests led him to NIIL, where his research studies focused on multimodal fusion for human behavioral and (psycho)physiological data. Along with Mohammad Naser, Namaz is a primary architect the SEMF framework described in Appendix Section D.2:

I see great benefit from [my role in] this project through exposure to the multimodal fusion, neuroscience, and psychophysiology in the context of real-world driving experiments. These are important complements to my previous experiences at social (macroeconomic, econometric) and theoretical (mathematical) scales. Ultimately, I want to leverage this experience into being a well-rounded human decision-making scientist.

Appendix C. Research Dissemination

This appendix highlights publications produced through the life of our project, which align with academic milestones, including thesis requirements and the mandatory requirement for student champions of each work to publish an article for graduation. Each of the seven summaries details the contributions of each publication, focusing less on individual data or results that were produced and more on the cohesive and progressive march toward the methodological solution initially envisioned for multimodal state classification.

This section is structured to highlight each individual's work, especially its student champions. The person pictured is always the first author, and all student names are fully spelled out and highlighted with bold and underlined text. For each publication, there is an overview page, followed by a summary, as well as all figures and tables from the student's original text. Our intention is to highlight the achievements of this team and display the extent and quality of technical outcomes from the program. The papers contained within this appendix represent significant added value; Appendix D is a technical supplement, delivering the remaining knowledge and engineering products.

Table C-1 summarizes the peer attention this work has garnered. By November 2024, these papers were cited 24 times, 100% by other groups,^{*} many of which were published in established and regularly cited journals.

Table C-1. Statistics for journals, papers, and people who have cited the works in this appendix.

Statistic	Journal	Paper median	Published authors
Average	55	13	34
Median	27	12	30
Maximum	169	26	67
Minimum	9	4	13

All values in Table C-1 are based on Google Scholar statistics, compiled through individual searches and examined in Microsoft Excel. The Journal column is based on Google's reported h5 values. These excluded four publications for which the metrics for the periodical or venue were not listed in the Google database,[†] as well as all pre-publication and thesis works ($n = 4$). Paper median represented the median value of the h-index for the author list of each paper; in cases where papers

^{*} This is noteworthy because the paper set predominantly contains papers that are preliminary; this kind of work typically gets self-citations before they are cited by others; and therefore, 100% of early, external citation activity for this number of papers stands out.

[†] These were as follows: 9th International Workshop on Advances in Sensors and Interfaces, IWASI 2023; Chaos, Fractals, and Complexity: COSA-Net 2022. Springer Proceedings in Complexity; AI Applications and Strategies in Teacher Education; and ICSTTE: International Conference on SmartRail, Traffic and Transportation Engineering.

had multiple authors with no prior publications, the number zero was only included once. Values in the Published authors column only included authors for whom an h-index was listed in Google or for whom could otherwise be estimated with their publication list using Google Scholar data. For completeness, Figure C-1 provides a distribution of all authors; indeed, there were a fair number of new authors ($h = 0$, $n = 13$) throughout the set of works cited. However, Table C-1 also demonstrates a degree of maturity in the remaining audience that has found the work suitable for citation in their own published papers.

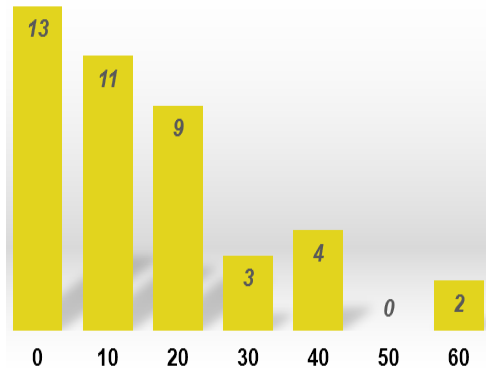


Figure C-1. Distribution of h-indices for all authors of published works cited in external papers in this appendix.

C.1 InfoSheets for Each Student-Led Paper

This section includes seven printable “infosheets” about the papers that were published during the project execution. Each infosheet summarizes a student-led publication and provides brief commentary about the work that is represented as it was situated in the overall project described in this report.

The first page of each infosheet includes a graphical header that provides the full bibliographic information for the published work, and this header is followed by an image of the first page of the published work and an excerpt from the abstract. Finally, the graphic on the bottom represents some information about the impact and acceptance of the paper since its publication, along with several bullet points highlighting the main contributions of each project.

The infosheets are followed by one or two more pages that provide an overview of tables, figures, and images included in each paper. The infosheets end with text that discusses the context of the work, the project-related challenge or challenge being addressed, new insights obtained, and, of course, future directions that were suggested.

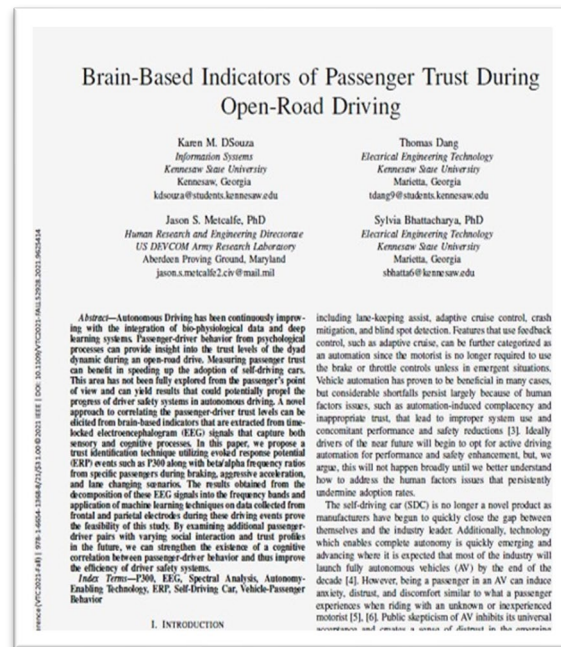


Student-Led Conference Paper

D'Souza KM, Dang T, Metcalfe JS, Bhattacharya S. Brain-Based Indicators of Passenger Trust During Open-Road Driving. IEEE 94th Vehicular Technology Conference (VTC2021-Fall); 2021. <https://doi.org/10.1109/VTC2021-FALL52928.2021.9625414>

Abstract (excerpt)

...Passenger-driver behavior from psychological processes can provide insight into the trust levels of the dyad dynamic during an open-road drive. Measuring passenger trust can benefit in speeding up the adoption of self-driving cars. This area has not been fully explored from the passenger's point of view and can yield results that could potentially propel the progress of driver safety systems in autonomous driving. A novel approach to correlating the passenger-driver trust levels can be elicited from brain-based indicators that are extracted from time locked electroencephalogram (EEG) signals that capture both sensory and cognitive processes. In this paper, we propose a trust identification technique utilizing evoked response potential (ERP) events such as P300 along with beta/alpha frequency ratios from specific passengers during braking, aggressive acceleration, and lane changing scenarios...



Keywords

P300, EEG, Spectral Analysis, Autonomy-Enabling Technology, ERP, Self-Driving Car, Vehicle-Passenger Behavior

Metrics

4 citations by other authors
16 median citing author h
48 maximum citing author h
169 maximum journal h5

Cited in DoD research

Highlights

- First successful publication from this project and first established and tested EEG method (P300 evaluation).
- Established the foundational approach for using natural driving events to explore human cognition during real-road driving.
- Pointed towards potentially detectable responses in evoked potentials and identified areas for immediate improvement.

In this paper, the authors explored the relationship between passenger trust and driver behavior through the lens of electroencephalogram (EEG) data, specifically focusing on evoked response potential (e.g., the P300 ERP), and beta/alpha frequency ratios during various driving maneuvers. The method relied on event-related analysis, wherein EEG signals were averaged across many standardized windows ($\pm 1/2$ s) around observable events in the driving records (such as acceleration, braking, and lane changes). By averaging across the standardized windows, the investigators could identify and further study systematic associations between brain activity and salient driving events. Finally, by then examining the same patterns in people who reported varying degrees of trust in the driver, the authors aimed to identify potential trust-related associations with brain dynamics in real-roadway driving. For this paper, the team produced a case study representing neutral versus high passenger trust. The reported results suggested a feasible path for using event-evoked brain activity for identifying trust levels in real-world circumstances, specifically by “locking” analyses to regularly occurring, salient, and meaningful driving events such as braking, lane changing, and aggressive acceleration. This study also revealed a gap that needs to be addressed for real-world applications of ERP-based approaches. Standard methods may have revealed non-stationary behavior in the brain data, which also implies specific directions that may be pursued in follow up.

Figure C-2 provides some high-level characteristics of the real-world driving (RWD) study, which covered an approximate 26-mile segment of U.S. Interstate Highway 95 in Harford County, Maryland.

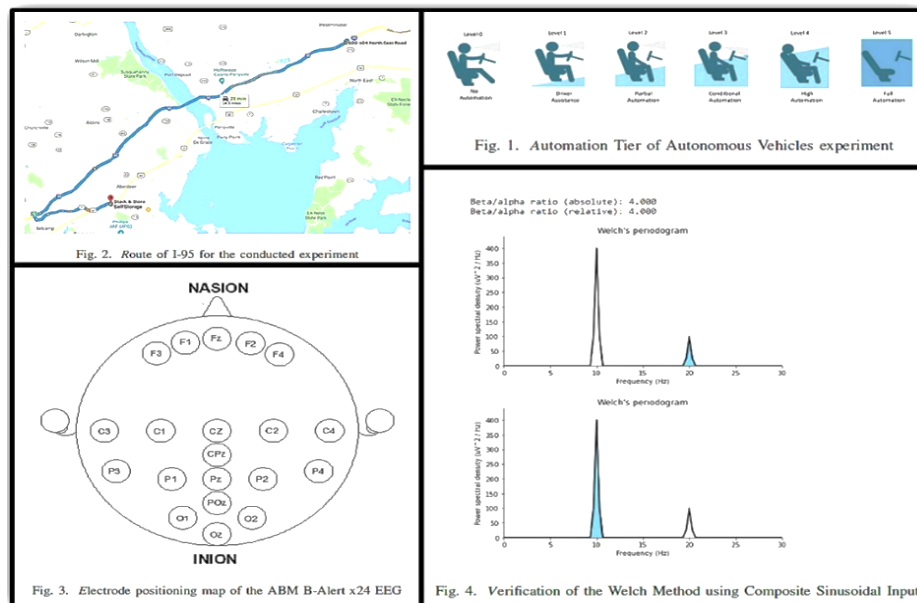


Figure C-2. RWD study information.

The data in Figure C-3, left middle and right top, show appropriate-magnitude evoked potentials among stable time series. The remaining charts show signals affected in detrimental ways. Successive efforts will deploy methods better at managing real-world data complexities like these signal drifts and variance heterogeneity. Successive analysis in joint time-frequency space, computed by wavelet transform, allowed greater insights into how and when unique signal variability manifested in meaningful ways.

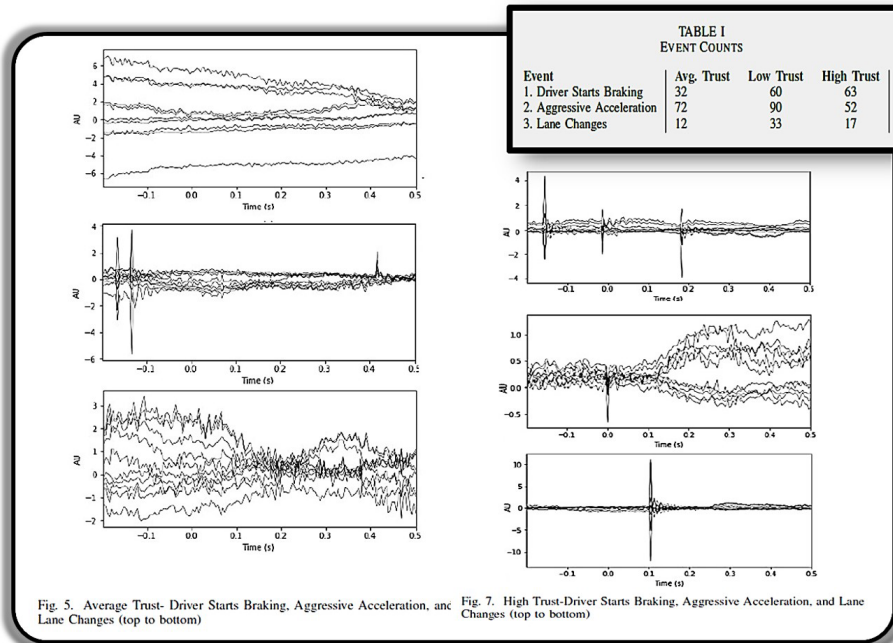


Figure C-3. Exemplar postprocessed EEG data recorded during the RWD study.

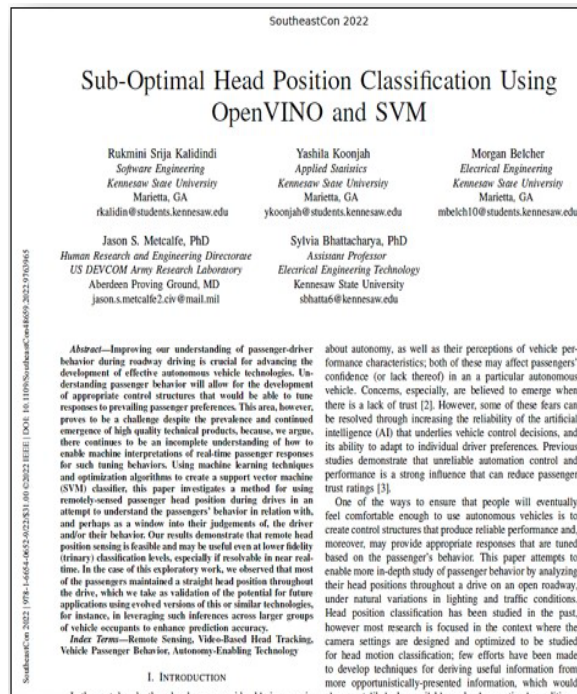


Student-Led Conference Paper

Kalidindi RS, Koonjah Y, Belcher M, Metcalfe JS, Bhattacharya S. Sub-optimal head position classification using OpenVINO and SVM. SoutheastCon; 2022. <https://doi.org/10.1109/SoutheastCon48659.2022.9763965>

Abstract (excerpt)

... Using machine learning techniques and optimization algorithms to create a support vector machine (SVM) classifier, this paper investigates a method for using remotely-sensed passenger head position during drives in an attempt to understand the passengers' behavior in relation with, and perhaps as a window into their judgements of, the driver and/or their behavior. Our results demonstrate that remote head position sensing is feasible and may be useful even at lower fidelity (trinary) classification levels, especially if resolvable in near real time. In the case of this exploratory work, we observed that most of the passengers maintained a straight head position throughout the drive, which we take as validation of the potential for future applications using evolved versions of this or similar technologies, for instance, in leveraging such inferences across larger groups of vehicle occupants to enhance prediction accuracy.



Keywords

Remote Sensing, Video-Based Head Tracking, Vehicle Passenger Behavior, Autonomy-Enabling Technology

Metrics

1st successful graduate thesis
1st complete subproject analysis with validation
80% average trinary classifier accuracy (**fourfold**)

Highlights

- Presents a valuable method for using suboptimal camera views for classifying occupant head position during open-road driving.
- Demonstrated integration of multiple ML methods (CNN, SVM) by creating a core feature for later application to the RWD data.
- Application of OpenVINO demonstrated an efficient and accessible approach to test and transition research products.

This work demonstrated how to recover head positions from video-based data recorded from suboptimal camera placements, specifically looking at human passengers in a moving car. Head data were captured by cameras mounted in unobtrusive locations to the side of either occupant's view, but this caused a suboptimal view frame. The initial goal of this was to combine head behaviors with other data modalities to infer responses to the behavior of the vehicle in traffic.

The main challenge was to analyze partial views of passengers' heads as an indicator of their relationship with the driver. The viewable proportion of the head was variable; for instance, as the passenger turned toward or away from the camera. The classification of head positions into left, right, and straight positions was achieved by a support vector machine ([SVM] a machine classification technique), which evaluated data features that were output from a convolutional neural network (CNN)-based image processor.

The research noted several challenges for this and similar methods. Classification accuracy varied, ranging from 50% to 98%. We concluded that there was natural variability in how each person moved their heads and this interacted with other environmental factors, such as ambient luminance and glare, to reduce the amount of classifiable imagery. This work provided a good model for integrating behavioral analytics into autonomous vehicle systems. Furthermore, by achieving these results with the open-source platform and established ML tools, the work demonstrated the relative ease with which value may be added in ways that were not as easily accessible and deployable in times past compared to today's world. Figure C-4 shows the in-cabin configuration on the left, whereas the right panel shows the trinary classifier labels for passenger head pose.

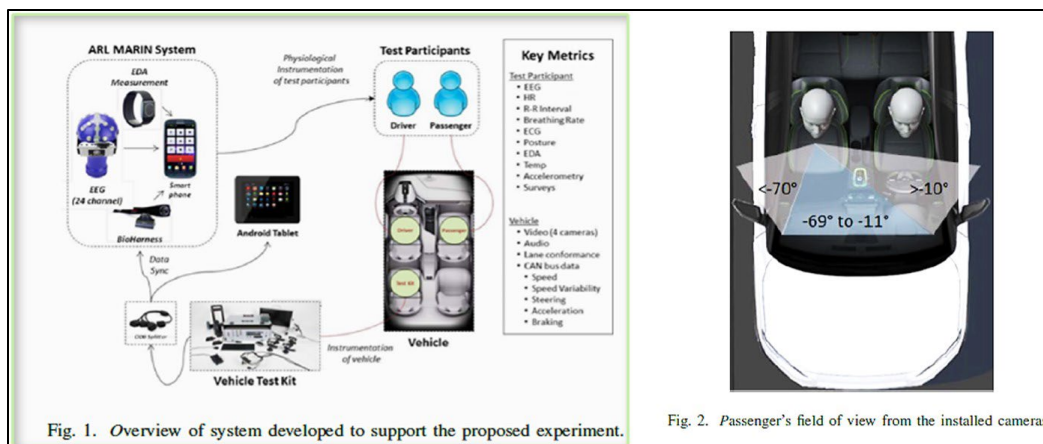


Figure C-4. Instrumentation and configuration for head pose classification during driving.

Figure C-5 is the complete diagram for this classifier system, and the charts in Figure C-6 provide tabular and graphic results for the head pose classifications (top

left and right, respectively), and classifier accuracy is shown from the fourfold cross-validations in the lower panel.

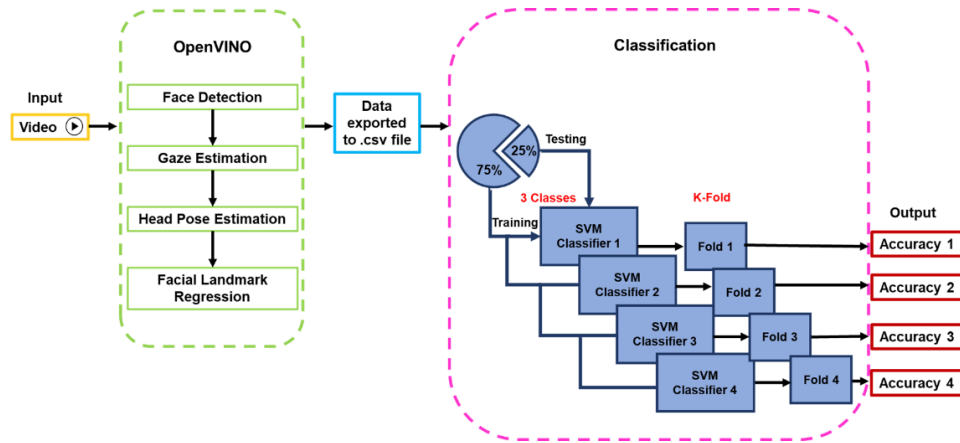


Figure C-5. Suboptimal head pose classification system.

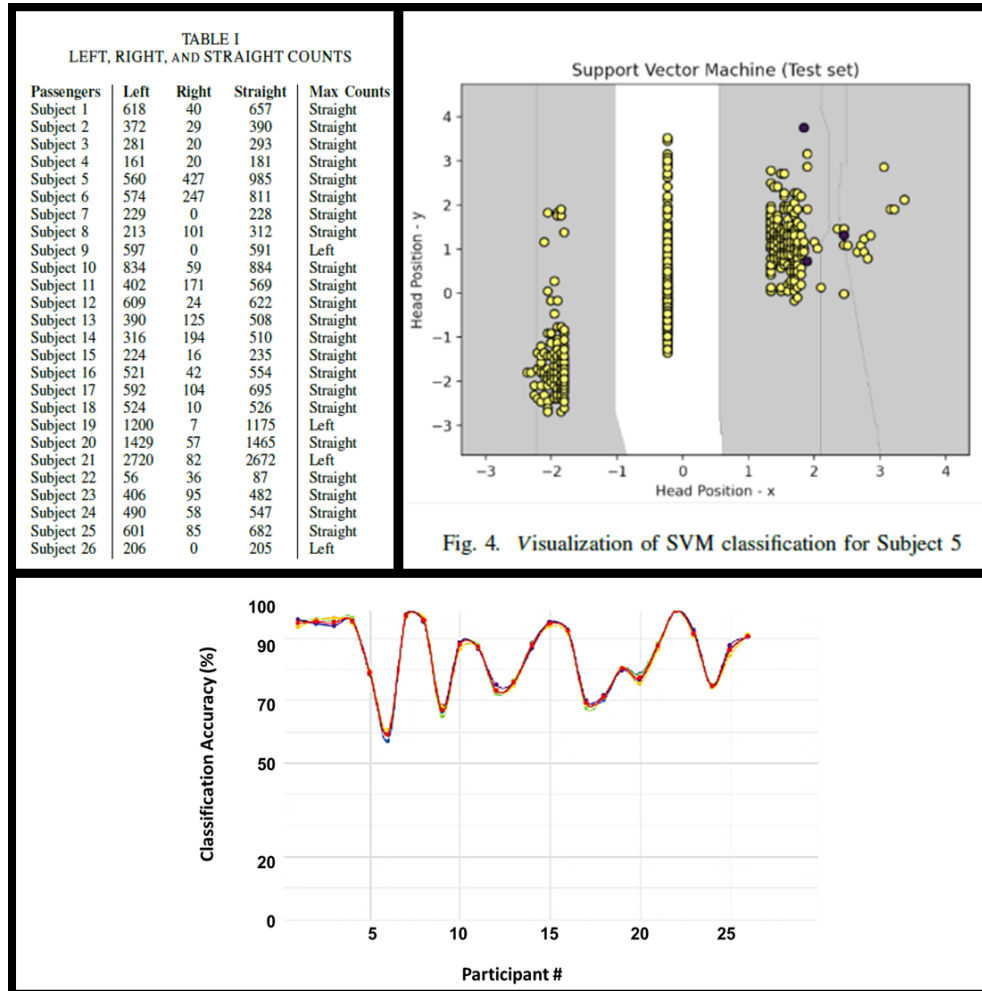


Figure C-6. Results and accuracy of classification applied to 26 passengers.

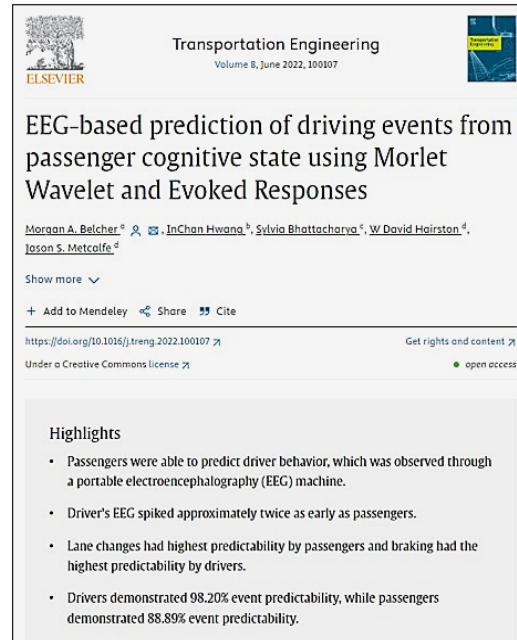


Student-Led Journal Article

Belcher MA, Hwang I, Bhattacharya S, Hairston WD, Metcalfe JS. EEG-based prediction of driving events from passenger cognitive state using Morlet wavelet and evoked responses. *Transp Eng.* 2022;8:100107. <https://doi.org/10.1016/j.treng.2022.100107>.

Abstract

This paper examines the predictability of driving events by passenger electroencephalography (EEG) data using Morlet wavelets and Evoked Responses (ER). As autonomous vehicles (AV) become more popular and the technology becomes more widely available, a major remaining challenge to overcome is passenger mistrust in the AV. People generally want to be in control of their vehicle, which is why human event predictability is such an important factor in AVs. This research study successfully predicted 93.61% of driving events including aggressive acceleration, braking, and lane changes on an open-road driving route.



Keywords

EEG, Autonomous Vehicles, Morlet Wavelets, Evoked Response

Metrics

14 citations by other authors
12 median citing author h
67 maximum citing author h
169 maximum citing journal h5

Cited in DoD research

Highlights

- Driver and passenger behaviors were detectable in individual trials using brain electrical activity (EEG).
- Lane changes appeared to be the most evidently detected or predicted by passengers, whereas braking was disproportionately detected by drivers.
- Suggests likely role-based differences in EEG features that will most predict external events of different types.

This paper explored the potential power of transforming the complex EEG data using an approach enabled by Morlet wavelets. Owing to a more precisely resolved time-frequency landscape, the wavelet approach is expected to yield more robust results with less favorable data signals. The objective of this approach was to evaluate signals of the driver and passenger during driving events, such as braking, aggressive acceleration, and lane changes. The report claims a high predictability rate of 93.61% in reference to the EEG signals reflecting impending driving events. This predictability refers to signal features that appeared in advance of the corresponding events but within only a short time window (<5 s). In this way, the features they identified with the Morlet wavelet were disproportionately ($>90\%$) predictive of relevant driving events. More importantly, as discussed in a corresponding report on the overall research program, this analysis revealed very high frequency features in brain activity that may have been indicative of passenger evaluations of the ongoing driving performance. Figure C-7 shows the percentages of activity spikes that were detected before the related driving events (Figure C-7, Table 1) and the average time between the spike and the event (Figure C-7, Table 2).

Table 1 Predictability of driving events.	
Category	Predictability (%)
Overall	93.61
Aggressive Acceleration	93.24
Driver - Aggressive Acceleration	97.30
Passenger - Aggressive Acceleration	88.89
Braking	91.89
Driver - Braking	100.00
Passenger - Braking	83.33
Lane Change	95.89
Driver - Lane Change	97.30
Passenger - Lane Change	94.44

Table 2 Average time (s) of spike before driving event.	
Category	Average Time (s) of EEG Spike
Driver - Aggressive Acceleration	-0.473
Passenger - Aggressive Acceleration	-0.268
Driver - Braking	-0.487
Passenger - Braking	-0.269
Driver - Lane Change	-0.496
Passenger - Lane Change	-0.253

Figure C-7. Results describing brain activity preceding driving events.

In the 2D plots shown in Figure C-8, signals demonstrate a qualitative stability with respect to their statistical properties compared with similar data processed using a Fourier Transform; here, baselines appear stabilized against significant linear drift, and the within-signal variance ranges are relatively homogenous across time and channel.

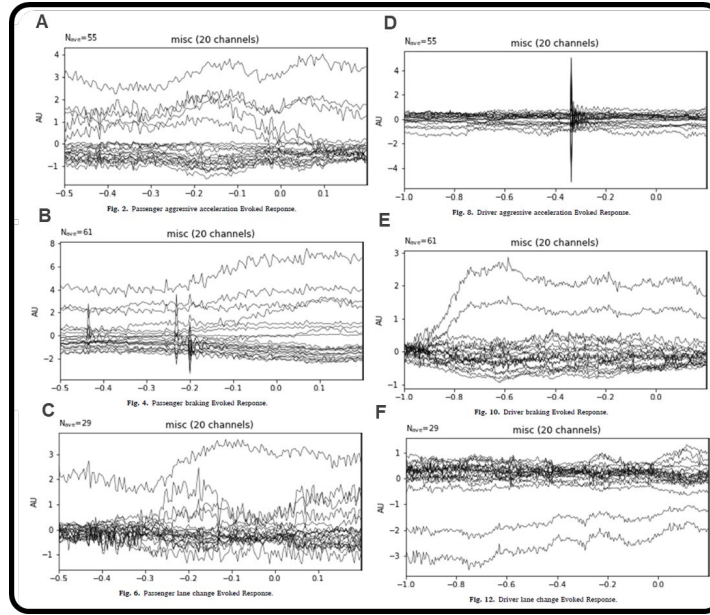


Figure C-8. Data from passengers and drivers, left and right columns, respectively, by driving event type (top to bottom: aggressive acceleration, braking, lane change).

Figure C-9 represents the same data as in Figure C-8, but it also shows the time-based fluctuations as a function of the frequency component. The letters within each chart will reveal appropriate comparisons between the two figure groups.

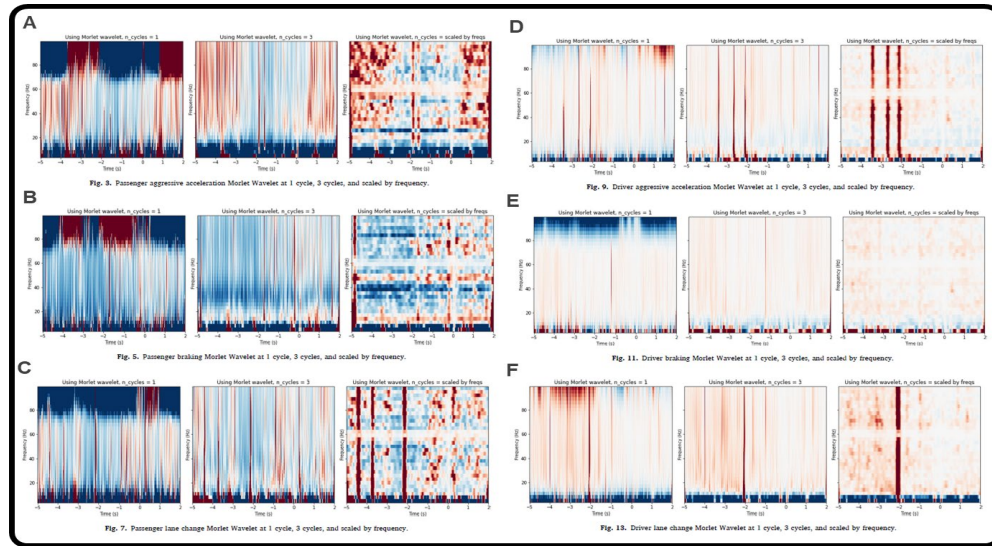


Figure C-9. Data from Figure C-8, decomposed across frequency within each time-step.



Student-Led Conference Paper

Reddy Dendi PK, Gurjala Y, Bhattacharya S, Metcalfe JS. A review study of smart vehicle seat sensor for real-time postural analysis. *Proceedings of International Conference on Computing and Communication Networks*; 2022 July. p. 229–241. DOI: 10.1007/978-981-19-0604-6_21

Abstract

The era of autonomous driving is gradually transitioning from humans as drivers to humans as passengers. With the advancement of artificial intelligence (AI), humans are switching roles from driver to passenger. Most of the early researches are driver central. But now it is essential to also study passenger data to understand human states in an autonomous vehicle as a passenger. Smart sensor monitoring systems can play a major role in assessing human and AI dynamics in an autonomous vehicle. Several research studies have utilized electroencephalography, electrocardiography, or electromyography technologies to study human states. But there is a huge gap in the practical applicability of such a sensor in everyday life. Hence, this paper conducts an extensive review of smart seat sensors for easy implementation in a car.

**Proceedings of International Conference on Computing and Communication Networks**

[Home](#) > [Proceedings of International Conference on Computing and Communication Networks](#) > [Conference](#)

A Review Study of Smart Vehicle Seat Sensor for Real-Time Postural Analysis

Praneeth Kumar Reddy Dendi , Yagna Gurjala, Sylvia Bhattacharya & Jason S. Metcalfe

Conference paper | First Online: 09 July 2022

591 Accesses | 1 Citations

Part of the book series: [Lecture Notes in Networks and Systems](#) (LNNS, volume 394)

Abstract

The era of autonomous driving is gradually transitioning from humans as drivers to humans as passengers. With the advancement of artificial intelligence (AI), humans are switching roles from driver to passenger. Most of the early researches are driver central. But now, it is essential to also study passenger data to understand human states in an autonomous vehicle as a passenger. Smart sensor monitoring systems can play a major role in assessing human and AI dynamics in an autonomous vehicle. Several research studies have utilized electroencephalography, electrocardiography, or electromyography technologies to study human states. But there is a huge gap in the practical applicability of such a sensor in everyday life. Hence, this paper conducts an extensive review of smart seat sensors for easy implementation in a car.

Keywords

[Driver monitoring system](#) [Smart cars](#) [Sensors](#) [Algorithm](#)

[Machine learning model](#)

Keywords

Driver Monitoring System, Smart Cars, Sensors, Algorithm, Machine Learning Model

Metrics

1 citation by other authors
1st student-led tech evaluation
90% observed accuracy across
36 reviewed works

Highlights

- Provides a considered overview of smart seat-based sensors for vehicle occupants.
- Suggests a potentially valuable roadmap for consideration of data from nonstandard, noninvasive sensor array.
- Offers point of challenge for our project: future analyses must consider and speak to applicability of successful methods.

This review focuses on the potential use of smart seat sensors for real-time postural analysis of passengers in roadway vehicles. Capturing postural behavior may provide an accessible window into the passenger psycho-emotional state and may enable a more robust and semantically contextualized means of sensing ongoing events both inside and outside the vehicle via body motion analysis. This paper evaluates an important challenge for the transition of research results about human state, which are normally derived from complex physiological measures that rely on sensors that have significant limitations in real-world circumstances.

Smart seat sensors are overlooked, yet potentially valuable technologies for improved passenger integration with automated vehicles. In particular, postural sensors are fairly robust across a significant range of force conditions, and postural behaviors are fundamentally associated with specific human states like drowsiness and fear. Furthermore, smart seat sensors may provide deeper insights into the passenger experience in ways that may usefully inform automated drivers of the future. Indeed, engineering challenges remain, such as assuring accuracy under conditions involving significant unexpected “biodynamic feedthrough,” where sensor readings are causally linked to complex vehicle acceleration instead of passenger states. However, modern approaches to signal processing and ML allow for effective removal of such effects. Considering the tightly constrained behavioral space (Figure C-10) that enables high prediction accuracies (Figure C-11), along with the potential for aggregation with other multimodal measures, we conclude that seat sensors provide an important opportunity for improving passenger experiences in automated vehicles.



Figure C-10. Illustration of behavioral limits in roadway vehicles under safety restraint (images generated with support of GPT 4o).

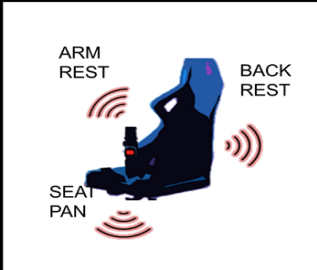
	Algorithm	Accuracy %	# Sensors	# Postures
	Support Vector Machine 1	97.0	Seat pan: 4	7
	Support Vector Machine 2	93.9	Seat pan: 8 x 8	9
	Random Forest	90.9	Arm rests: 2 Back rest: 4 Seat pan: 10	7
	AdaBoost	87.0	Seat pan: 10	15
	Naïve Bayes	82.0	Seat pan: 6 x 8 Back rest: 2 x 8	9

Figure C-11. Comparison of several classification methods applied to smart seat sensors.



Student-Led Conference Paper

Biswas S, Hairston WD, Metcalfe JS, Bhattacharya S. EEG based BCI for autonomous control: a review. SoutheastCon; 2023 April. DOI: 10.1109/SoutheastCon51012.2023.10115194

Abstract

Brain-computer interface (BCI) has had exciting improvements in the last decade due to the developments in brain science and computer technology, making BCI a top research area in applied science. With the advancement of technologies, researchers are now experimenting using BCI for operating vehicles, robots, and aircraft. However, choosing the suitable brain signal acquisition technique among the available ones is a big concern. In this paper, we will survey the papers that used the BCI framework for controlling vehicles and quadcopters, its limitations, its prospect, and the reason electroencephalogram (EEG) best suits the BCI framework.

EEG based BCI for Autonomous Control: A Review

Susant Biswas
Dept. of Analytics and Data science
Kennesaw State University
Marietta, 30060, GA, U.S.
sbiswas1@students.kennesaw.edu

Dr. W David Hairston
US DRYCOM Army Research Laboratory
Aberdeen Proving Ground, MD, U.S.
MD, U.S.
william.d.hairston@army.mil

Dr. Jason S. Metcalfe
US DRYCOM Army Research Laboratory
Aberdeen Proving Ground, MD, U.S.
MD, U.S.
jason.s.metcalfe@army.mil

Dr. Sylvia Bhattacharya
Dept. of Engineering Technology
Kennesaw State University
Marietta, 30060, GA, U.S.
sbhattach@kennesaw.edu

Abstract—Brain-computer interface (BCI) has had exciting improvements in the last decade due to the developments in brain science and computer technology, making BCI a top research area in applied science. With the advancement of technologies, researchers are now experimenting using BCI for operating vehicles, robots, and aircraft. However, choosing the suitable brain signal acquisition technique among the available ones is a big concern. In this paper, we will survey the studies that used the BCI framework for controlling vehicles and quadcopters, its limitations, its prospect, and the reason electroencephalogram (EEG) best suits the BCI framework.

Index Terms—Brain-computer interface, electroencephalogram (EEG), brain science and computer technology, signal acquisition technique, vehicles and quadcopters

I. INTRODUCTION

The concept of controlling prosthetic limbs using brain signals was first explored in the 1970s. Since then, scientists have been working to decipher brain wave patterns to provide more precise and practical control over external devices. And a Brain-Computer Interface (BCI) is a tool that permits neural activity in the brain to remotely interact with people or control robots, artificial limbs, or machinery [1]. [2] No external devices or muscular activity is needed to actuate commands and complete them [3]. The research community initially developed BCI for biomedical applications, leading to assistive device generation [4]. They have made it easier for physically challenged people to regain their capacity to move around and replace lost motor functionality [5]. The scientific community has been inspired to investigate how BCI can be employed in other areas of science and technology.

Noninvasive neuroimaging techniques can be used to capture brain activity [2], [6], yet EEG is the most widely used for the BCI framework. EEG detects electrical activity in the brain, particularly in the cortex but also in deep brain areas, that results from the flow of electric currents during synaptic activation of neuronal dendrites. Electrodes record the electric impulses from the scalp. Wheelchairs and communication aid systems have both been controlled by EEG signals in the past [7]. The control of assistive and rehabilitative equipment

has also benefited from the development of EEG approaches during the past decade [8].

Recently, many researchers have concentrated on real-space BCI-controlled vehicles and flying robot navigation. Successful implementation of the technique in the area will benefit the physically challenged as well as general people providing a lot of feasibility while operating them. With the development of the research, a review of the works and proper guideline is also required.

The objective of this work is two-fold. First, we aim to review current EEG based BCI technologies; second, their applications to vehicle and flying quadcopter operations, primarily in military applications. The rest of the paper is structured as follows. Section II contains an overview of the most up-to-date brain signal acquisition technologies, along with their advantages and limitations. Section III will discuss why EEG is best suited for BCI in most cases and their applications to vehicles and flying robots. Section IV will cover where we can apply EEG-based BCI in military applications. And section V concludes the paper with future direction.

II. BRAIN SIGNAL DATA ACQUISITION

The techniques used for neuroimaging can be split into two categories: invasive and noninvasive. Invasive BCI involves placing electrodes into the brain. Non-invasive BCI reads brain impulses using sensors or medical scanning equipment installed on caps or headbands. That is, non-invasive BCI does not require surgery, whereas invasive BCI requires installing sensors inside the brain. Table I

Electroencephalography (EEG) and electrocorticography (ECoG) are the most prevalent invasive and noninvasive technologies, respectively. ECoG is recorded from the cortical surface, but other invasive technologies record signals from within the brain. The high-quality spatial and temporal characteristics of invasive signals lead to the successful decoding of homeostatic parameters [9]–[11]. The risk of surgery and the recorded data's progressive deterioration are major disadvantages of noninvasive electrodes since they collect data from within the brain. That led to an increase in the

Keywords

Brain–Computer Interface, Electroencephalogram (EEG), Brain Science and Computer Technology, Signal Acquisition Technique, Vehicles and Quadcopters

Metrics

3 citations by other authors
12 median citing author h
48 maximum citing author h
26 maximum journal h5

Cited in DoD research

Highlights

- The breadth of this review provides an indication of use-cases for BCI and occupant sensing for improving safety.
- Offers insights and suggestions as to modern ways that BCI may still apply more broadly than driving, such as quad rotor-based drones and military applications.

This paper provides an overview of advancements in portable BCI technology, focusing on EEG-based BCI for ground and air vehicle systems. BCI technology may still provide revolutionary changes in control systems for various applications, extending beyond medical devices to include vehicles, robots, and small uncrewed aircraft. A focal aspect of the review is the discussion of selecting brain-signal acquisition techniques, with a special emphasis on the suitability of EEG for BCI frameworks. Particularly successful applications of EEG signals were noted for interpreting user intentions, which has applicability to formulating sparse control commands to produce real-world outcomes. Moreover, it is trivial to forecast that intelligent BCI systems may be further enabled with modern ML methods, including SVM, CNN, and ensemble classifiers.

There is a critical need to address practicality, reliability, and commercial viability of these BCI for any sort of real-world, safety-critical operation. Images in Figure C-12 represent the fundamental difficulty of limiting “BCI” applications to brain measurement only. That is, the majority of methods and sensors for obtaining spatially and temporally precise records of brain activity are generally large, intrusive or obstructive, non-portable, and preclude application in any typical circumstances. The tables provided in Figure C-13 represent this challenge similarly, but from a perspective of acquisition technique as opposed to hardware; again, this reinforces the considerable challenge of translating brain-measurement techniques into real-time BCI applications in actual operational tasks.

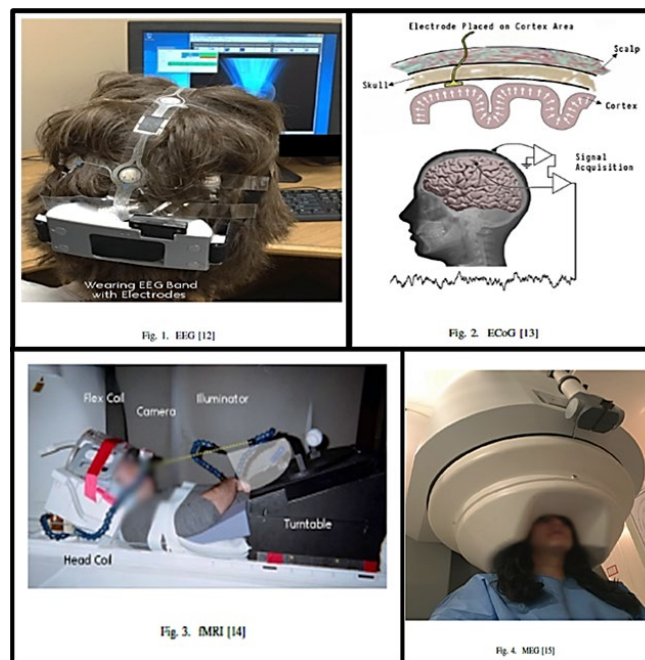
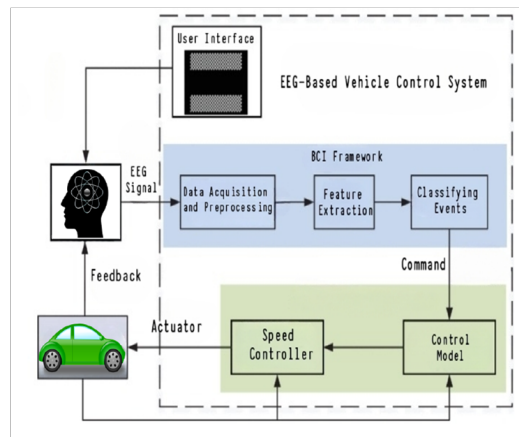


Figure C-12. Examples of common methods for observing and recording brain activity.

Method	Description	Advantages	Limitations
Electrocorticography (ECoG)	Electrodes implanted on surface of cerebral cortex, under the skull bone)	✓ High spatial and temporal resolution ✓ Better signal-to-noise ratio	✗ Surgical procedure required
Electroencephalography (EEG)	Electrodes places on skin of the external scalp	✓ High temporal resolution ✓ Affordable, non-invasive, and portable	✗ Low spatial resolution ✗ Difficult to localize source of activity
Functional Magnetic Resonance Imaging (fMRI)	Magnetic sensors detect changes in regional blood flow of scalp	✓ High spatial resolution ✓ Identifies active brain regions	✗ Expensive ✗ Requires the subject to remain still
Magnetoencephalography (MEG)	Detects magnetic fields within the brain from an external sensor array	✓ Very high spatial and temporal resolution ✓ Accurate source localization	✗ Sensitive to magnetic interference ✗ Expensive



Types	Details
Invasive	Electrodes are implanted directly onto the brain, requiring surgery. Examples include electrocorticography, single neuron action potentials, multi-unit activity, and local field potentials.
Non-Invasive	Sensors or devices are mounted externally (e.g., caps or headbands) to record brain signals without surgery. Examples include electroencephalography, functional magnetic resonance imaging, magnetoencephalography, and near-infrared spectroscopy.

Figure C-13. Overview of traditionally applied BCI and brain-measurement methods.

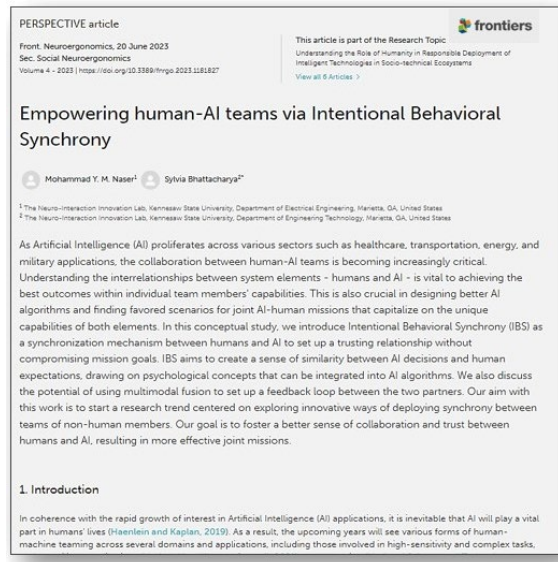


Student-Led Journal Article

Naser MYM, Bhattacharya S. Empowering human-AI teams via intentional behavioral synchrony. Front Neuroergon. 2023;4. <https://doi.org/10.3389/fnrgo.2023.1181827>

Abstract (*excerpt*)

... In this conceptual study, we introduce Intentional Behavioral Synchrony (IBS) as a synchronization mechanism between humans and AI to set up a trusting relationship without compromising mission goals. IBS aims to create a sense of similarity between AI decisions and human expectations, drawing on psychological concepts that can be integrated into AI algorithms. We also discuss the potential of using multimodal fusion to set up a feedback loop between the two partners. Our aim with this work is to start a research trend centered on exploring innovative ways of deploying synchrony between teams of non-human members. Our goal is to foster a better sense of collaboration and trust between humans and AI, resulting in more effective joint missions.



Keywords

Human–AI Teaming, Intentional Behavioral Synchrony (IBS), Human–Machine Trust, Human–Machine Interaction (HMI), Multimodal Fusion

Metrics

2 citations by other authors

1st student-developed theoretical construct

>1,100 online views

>370 downloads

Highlights

- Presents a vision of using a natural human attraction towards behavioral synchrony to improve alignment in human–AI teams.
- Examines the generality of the proposed construct in relation to domains of psychology, BCI, and engineering.
- Offers a vision for AI schemes that do not converge to a single path or solution, but maintain an array to more flexible support partnering with people of diverse priorities and expectations.

In this conceptual paper, Naser and Bhattacharya explore the notion of *Intentional Behavioral Synchrony* (IBS) as a mechanism to foster trust and enhance collaboration within human–AI teams. Drawing from psychological insights, they propose that by aligning AI decision-making patterns with human expectations in a manner that reflects human-like decision processes, it is possible to establish a foundation of trust without compromising the objectives of a mission. Such an alignment may derive from a partnership and collaborative dynamic. The concept of IBS underscores the potential to adaptively operate across interactions in a personalized manner, accounting for individual differences in human–AI interaction dynamics. The potential of using multimodal fusion techniques to establish a feedback loop is a core suggestion, which will enhance responsiveness to human signals that are often manifest in multiple contextualized events. The concept of behavioral synchronization and its power for regulating coordination in complex systems is well founded in scientific theory across domains as disparate as non-equilibrium thermodynamics and infant cognitive development. It offers a promising avenue for developing AI systems that can better align with human expectations and behaviors, which is also an agreeable suggestion.

Figure C-14 depicts a complex, layered, and hierarchical situation for behavioral synchrony. With agents, learning is driven toward optimality via typical rewards. These paths risk fixity, thus limiting value for generalized cooperation in heterogeneous teams. To preserve cooperation opportunities, agents ought to prioritize solutions that optimize or satisfice across multiple outcomes. Singular agents maintain many decision and action pathways, and thus they are envisioned to better subserve the interests and nuanced teamwork in a complex collective.

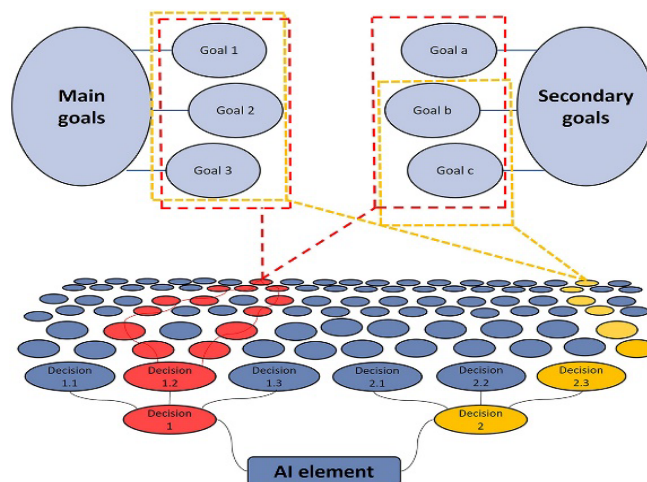
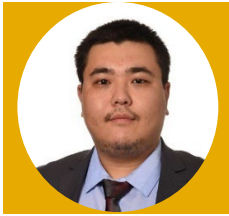


Figure C-14. Fixed vs. flexible solution paths in complex goal spaces for decision agents.



Student-Led Conference Paper

Ishmakhametov N., Naser MYM, Bhattacharya S. Assessing the rigor of machine learning in physiological signal processing applications. IEEE SoutheasCon; 2024.
<https://doi.org/10.1109/SoutheastCon52093.2024.10500274>

Abstract (excerpt)

... This review analyzes 82 influential articles (2018-2023) from IEEE Xplore, aiming to identify weaknesses and assess overall rigor. It emphasizes the need for enhanced research quality and reproducibility. The key findings reveal that in over half of the articles, the ratio of female-to-male participants recruited for data collection is below 50%. Additionally, nearly 30% of the studies involve fewer than 10 subjects in data collection, with only 7% providing justification for their sample size. Moreover, only about 34% of the articles provide access to their data, and a mere 26% report performance using a confusion matrix. These insights underscore critical areas for improvement, enhancing the robustness and transparency of applications in the physiological signal processing domain.

Keywords

Machine Learning, Physiological Signals, Rigor, Replicability, Reproducibility, Synchronization, EEG, EMG, ECG

Assessing the Rigor of Machine Learning in Physiological Signal Processing Applications

Publisher: IEEE

[Cite This](#)

[PDF](#)

Namazbai Ishmakhametov ; Mohammad Y.M. Naser ; Sylvia Bhattacharya [All Authors](#)

32

Full

[Text Views](#)



Abstract

Document Sections

I. Introduction

II. Methods

III. Results

IV. Discussion

Authors

Figures

Abstract:

In the dynamic field of biomedical engineering, the pervasive integration of machine learning into physiological signal processing serves various purposes, from diagnostics to Brain-Computer Interface (BCI) and Human-Machine Interface (HMI) using techniques such as Electroencephalography (EEG), Electromyography (EMG), Electrocardiography (ECG), and others. Nonetheless, the inherent scientific diversity within biomedical research often poses challenges, with practices sometimes misaligned with machine learning and standard statistical principles. This review analyzes 82 influential articles (2018–2023) from IEEE Xplore, aiming to identify weaknesses and assess overall rigor. It emphasizes the need for enhanced research quality and reproducibility. The key findings reveal that in over half of the articles, the ratio of female-to-male participants recruited for data collection is below 50%. Additionally, nearly 30% of the studies involve fewer than 10 subjects in data collection, with only 7% providing justification for their sample size. Moreover, only about 34% of the articles provide access to their data, and a mere 26% report performance using a confusion matrix. These insights underscore critical areas for improvement, enhancing the robustness and transparency of applications in the physiological signal processing domain.

Metrics

1st NIIL PRISMA-based review

82 influential articles included

15 rigor criteria suggested

2/3 of reviewed works failed to address rigor criteria

Highlights

- Provides a detailed review of current state of rigor in ML research involving human physiological signals.
- Uniquely reviews biomedical applications of ML.
- Identifies a clear need for improved attention to reporting of empirical details in ML research.

In this paper, Ishmakhametov et al. explore the level of rigor of the most influential IEEE papers published in journals in the last 5 years, focusing on works that incorporated physiological signals with classic and deep learning approaches. Ishmakhametov et al. offered a practical five-step strategy to examine the rigor of any ML work involving human physiological data. First, they evaluate the soundness of the research question, and then they build a physiological data strategy mindful of quality concerns, and finally they ensure replicability by providing access to data and scripts. The review also highlights that only 5 of 65 highly cited IEEE papers satisfied each rigor criteria.

Figure C-15 illustrates the recommended procedure for systematic evaluation of research in physiological ML. Shown as a top-down flow (Figure C-15, top left), the questions in Table 1 correspond with flowchart items B–E. Table 2 (Figure C-15, lower left) presents a final assessment of the 82 reviewed articles, per the suggested criteria.

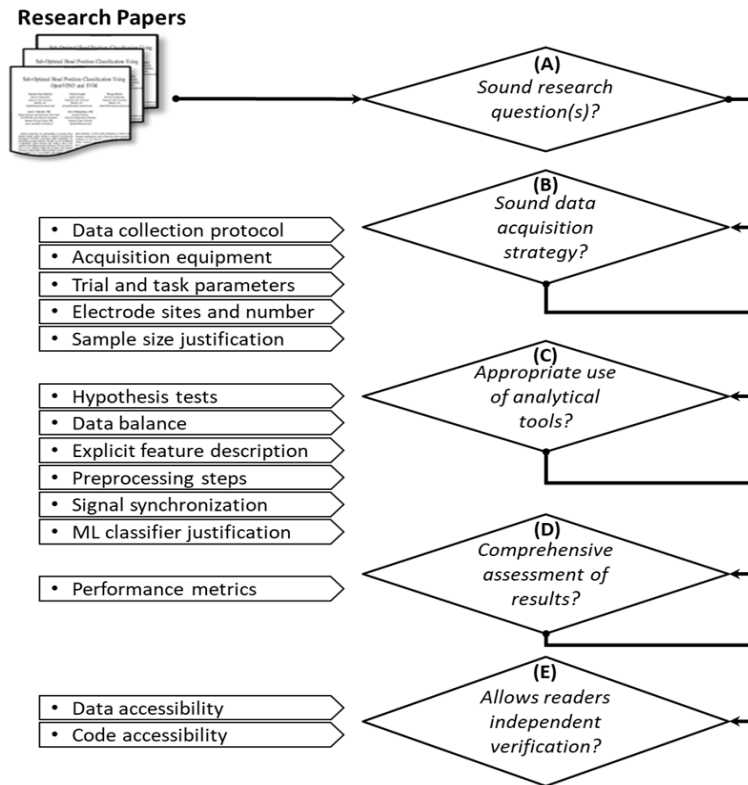


Figure C-15. Summary of the proposed method for evaluation of the rigor or physiological ML research reports.

Using the well-established PRISMA* methodology for systematic literature review, Ishmakhametov et al. identified 238 papers from the IEEE eXplore database and successively filtered the initial list down to 82 based the stepwise application of a predefined set of quality criteria, assuring that the body of literature selected could be evaluated as a collective against clearly stated standards.

This systematic approach allowed quantitative evaluation of whether and to what extent the group of papers reflected a research community that was attentive to these standards of rigor. For example, Figure C-16 reveals the heavy degree to which physiological ML research is based on relatively small sample sizes (left panel) and appears significantly imbalanced toward an overrepresentation of data obtained from biologically male as opposed to biologically female participants.

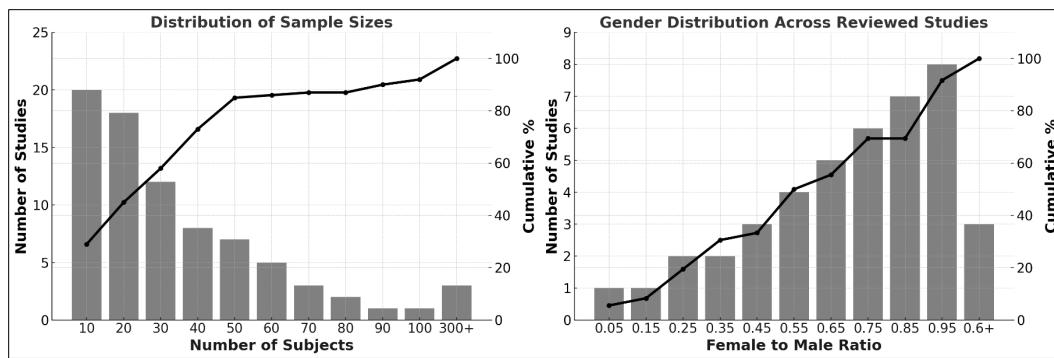


Figure C-16. Charts representing the reported use of relatively small samples that are imbalanced toward biologically male-heavy study populations.

A well-known challenge in the research literature involving biomedical monitoring technologies, which often rely on applications of physiological ML, is the inherent limits it places on the representative quality of the data to be obtained both in terms of sample size and demographic balances. In this report, the primary value of this exercise was in highlighting areas where our own research must address and improve the level of rigor throughout design. Based on this work, we designed and conducted a well-received research reproducibility workshop during the IEEE SoutheastCon 2024 annual conference.

* For details, see PRISMA for systematic review protocols [PRISMA-P], Transparent Reporting of Systematic Reviews and Meta-Analysis. <http://www.prismastatement.org/Extensions/Protocols>

Appendix D. Technical Supplement

This appendix provides an archive of the final state of the overall effort in terms of the final technical deliverables (summary in Table D-1). “The long-range objective of this research was to develop a method to use a multimodal set of sensors, including physiological, behavioral, and environmental, to enable reliable inference of human states and cognition during interactions with a vehicle-based driving autonomy” (as noted in Section 2.2).

Table D-1. Technical deliverables: data, documentation, and software.

Deliverable no.	Appendix subtitle	Contents
D.1	Passenger Experience Data	Descriptive metrics for each passenger experience
D.2	Multimodal Fusion and Human State Estimation	Progress report on multimodal method development and immediate path forward
D.3	Code Catalog	Descriptions of the open ARL-KSU GitHub repository contents

D.1 Data Table: Descriptive Metrics for Each Passenger Highway Experience

Table D-2 summarizes the experiences of 39 of 43 passengers in the U.S. Army Combat Capabilities Development Command Army Research Laboratory Real-World Driving (RWD) experiment, framed in the context of the experience of driving events. These data are intended to be a unique depiction, and they provide a touchpoint regarding the basic experiential aspects of *being a vehicle passenger* on open, improved roadways, in a developed region (U.S. Mid-Atlantic), with circa 2000 C.E. vehicle technology.*

* To complete the utility of these data, we provide the following geographic and date information related to the exact traffic context of the period of data collection. Route spanned a segment of U.S. Interstate 95 between exit numbers 80 and 100 at roadways MD 543 (“Riverside Parkway”) and MD 272 (“North East Road”). All data were obtained between November 2017 and August 2018.

Table D-2. Summary of passenger highway experience.

	Total drive duration (min)	Highway duration (min)	Max. highway speed (mph)	Avg. highway speed (mph)	% highway driven >65 mph	Aggressive accelerate (no.)	Braking (no.)	Lane changes (no.)	Turn signal use (no.)	Lane variance (cm)	Lane warnings (no.)	Headway warnings (no.)
1	58.30	35.57	71	60.2	38	31	24	13	17	11.4	...	24
2	59.62	37.79	70	58.31	13	31	23	5	14	10.6	2	24
3	58.60	34.46	76	65.4	76	32	20	41	41	18.4	...	34
4	49.18	38.24	68	52.0	7	23	21	8	16	10.9	...	10
5	58.60	34.46	76	65.4	76	32	20	41	41	18.4	...	34
6	62.88	35.3	76	63.8	73	35	34	25	33	17.7	...	28
7	59.45	38.85	71	53.2	10	29	18	9	9	11.0	2	2
8	55.98	34.92	73	62.7	47	31	26	14	16	25.1	...	28
9	57.52	37.04	69	57.4	10	23	33	14	21	15.0	...	10
10	54.85	35.46	71	61.0	30	35	22	21	26	16.1	...	20
11	58.76	34.99	75	62.8	53	18	33	22	23	16.7	2	32
12	60.57	37.07	70	55.4	13	17	21	7	12	12.7	...	22
13	58.46	36.48	70	58.1	17	25	21	9	10	12.7	...	14
14	58.89	37.88	68	51.1	2	30	32	13	15	12.5	...	6
15	59.00	36.98	71	59.6	33	27	36	14	17	16.9	2	26
16	70.32	36.85	66	56.6	3	17	19	16	21	24.1	...	14
17	56.07	33.95	75	64.8	64	26	33	23	22	14.1	...	14
18	58.8	35.5	71	60.6	36	28	23	12	15	14.0	2	26
19	55.08	36.58	66	55.0	8	28	22	9	20	8.4	...	8
20	70.63	37.30	66	55.6	3	27	32	11	17	12.1	...	4
21	78.58	52.27	73	61.5	29	26	15	41	38	21.6	2	20
22	56.21	34.66	69	61.4	37	24	10	28	31	23.0	...	26
23	59.73	37.02	71	57.5	20	28	37	23	28	21.6	...	24
24	65.75	36.84	69	57.6	7	31	22	23	30	19.1	1	10
25	85.57	38.4	66	53.2	1	30	26	17	23	22.9	1	10
26	64.46	38.19	72	57.7	42	36	60	18	19	17.3	3	28
27	55.07	35.35	70	59.2	50	21	30	20	22	46.9	1	8
28	90.36	37.2	71	56.6	9	29	41	11	13	12.7	...	28
29	60.69	38.74	67	53.3	2	32	18	15	22	16.7	...	2
30	62.97	35.25	75	62.3	45	28	10	21	23	17.8	...	18
31	62.03	40.73	67	48.2	2	40	78	5	12	21.8	2	48
32	64.66	37.84	70	55.2	17	28	41	24	28	49.1	...	14
33	59.03	37.46	71	56.7	18	24	69	13	20	57.8	1	2
34	59.70	38.53	70	55.0	15	19	13	13	19	12.8	...	16
35	64.02	37.17	68	57.6	13	15	32	17	11	22.9	2	24
36	59.93	36.67	71	56.1	14	25	17	9	14	13.6	...	10
37	59.91	37.65	64	53.0	0	23	12	10	15	8.6	...	22
38	61.10	38.17	66	52.9	7	27	21	7	10	7.0	...	4
39	55.71	34.13	76	64.2	54	36	19	17	18	11.9	2	36

D.2 Multimodal Fusion and Human State Classification Progress Update 2024

By September 2022, the Kennesaw State University (KSU) team had been producing and refining software to extract quantitative features from the biophysical data (electroencephalography [EEG; the method of recording and monitoring brain electrical activity from surface electrodes on the scalp], electrocardiography [ECG; the method of recording and monitoring heart electrical activity from surface sensors placed on the body], and Head Pose) that were obtained during DEVCOM Army Research Laboratory's RWD experiment. These features would form the basis of the multimodal classification of external driving events from the human-sensed data in the spirit of the "human as a sensor" concept. Indeed, the first year demonstrated for the team that there would be some formidable challenges, such as potentially nonstationary properties of the EEG signals,¹ but also happy successes in progressing rapidly deployable methods for recovering features from data previously thought to be too incomplete to be usable.² With some students graduating and others adopting positions of leadership, the team was encouraged by the progress in addressing the real-world EEG data.³

Table D-3 lists all features from the RWD dataset that have been explored as of this report. Of particular importance for this project was the question of how to address the considerable potential of statistical distortions, owing to likely collinearities existing among the redundant measurements taken on the same participant (e.g., shared variance, or common information, affecting changes and patterns in multiple variables), and that these collinearities would compromise the ability to evaluate true associations that exist among the psycho-physiological modalities. Therefore, before proceeding with developing any multimodal approach, we assessed whether the originally anticipated improvements from multimodal measurement would necessarily manifest among the varied layers of complexity associated with roughly equal potential for distortion of results.

¹ D'Souza KM, Dang T, Metcalfe JS, Bhattacharya S. Brain-based indicators of passenger trust during open-road driving. 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall); 2021. <https://doi.org/10.1109/VTC2021-FALL52928.2021.9625414>

² Kalidindi RS, Koonjah Y, Belcher M, Metcalfe JS, Bhattacharya S. Sub-optimal head position classification using OpenVINO and SVM. SoutheastCon 2022; 2022. <https://doi.org/10.1109/SoutheastCon48659.2022.9763965>

³ Belcher MA, Hwang I, Bhattacharya S, Hairston WD, Metcalfe JS. EEG-based prediction of driving events from passenger cognitive state using Morlet wavelet and evoked responses. Trans Eng. 2022;8:100107. <https://doi.org/10.1016/j.treng.2022.100107>

Table D-3. Passenger data features for real-world event classification.

Modality	Data Feature	Estimator
EEG		
20-channel, full-scalp distribution; Pre-processed with ARL ABDHS Pipeline; Estimates used 2-s, event-centered epochs	Alpha	Power density, 8- to 12-Hz interval
	Beta, low	Power density, 13- to 15-Hz interval
	Beta, mid ^a	Power density, 15- to 20-Hz interval
	Beta, high	Power density, 20- to 28-Hz interval
	Gamma, low	Power density, 40- to 50-Hz interval
	Gamma, high	Power density, 70+ Hz interval
ECG		
Single lead, 2-point, chest-borne ECG system; Processed with HeartPy toolkit for Python; Estimates based on 5-s epochs	Interbeat interval	Avg. peak-to-peak (R-R) interval, ms
	Heart rate	Number of peaks per observation period, bpm
	Rate variability	Std. dev., normal-normal heart period, ms
	Interval variability	Root-mean-square of successive (interbeat) differences, ms
	Short-term interval variability	Std. dev. of successive (interbeat) differences
	Short-interval influence ^a	Proportion of interbeat differences >20 ms
Eye		
Head pose was a proxy for gaze direction; Estimated with 10-s video epochs	Dwell time left	Duration of persistence of support vector machine (SVM) class ID in left region, seconds
	Dwell time center	Duration of persistence of SVM class ID in center region, seconds
	Dwell time right	Duration of persistence of SVM class ID in right region, seconds
Seat		
Encoded outputs from each sensor in a 2 × 3 array of force-sensing resistors (FSRs); Estimated in 2-s epochs	Backward pressure distribution	Avg., std. dev. of left, center, and right rear FSR pads
	Forward pressure distribution	Avg., std. dev. of left, center, and right front FSR pads

^a Features that were excluded from early (unimodal and multiclass) simulations.

D.2.1 The State of the Art in Multimodal Fusion: Observations from Literature Review

Without clear conceptual guidance, the question is not just whether but also when to fuse extracted data features.* Early fusion occurs when features are aggregated before being subject to the classification and cross-validation processes (Table D-3). Such approaches appear to gain advantage through weight of information (amount of correlated data) applied to the problem. However, in early fusion approaches, there is also the likelihood of complex aggregation and amplification of noise and other sources of distortion due to these same correlations. Late fusion approaches thus perform classifications more specifically within

* This section and the proposed Sequential, Event-based Multimodal Fusion approach were co-developed and authored by Mr. Mohammad Y. M. Naser and Namazbai Ishmakhametov. Section edits and text additions were incorporated by first author (JSM) to maintain an integrated voice and narrative tone as throughout the remainder of the documents here.

modalities first, where the variance structures may be precisely modeled without cross-modality collinearities, and then fusion is applied at the classifier output layer, as if different groups “voted” in an electoral process; such methods are also known as ensemble models.

While preparing our literature review, we focused on this and other quality-assessment issues with machine fusion and classification with multiple biomedical data modalities.⁴ We also observed a tendency for works that used early fusion to offer less than a 5% increase in accuracy. However, for those using late fusion, the same body of literature appeared to show greater improvement, but they still lingered below 10% in the aggregate. Considering known issues with biases in published works toward positive results—for example, inflating effect sizes relative to what might be expected in most cases—these findings amplified our caution regarding our attempts to optimize multimodal methods; that is, we decided to examine how much we gained from multimodal methods when applications of unimodal approaches may suffice and be less costly.

Even more foundational from the works reviewed, it appeared that most biomedical multimodal research focused on identifying optimal modality combinations without fully considering the objective function. For instance, the objective function, as defined by the event to be classified, has intrinsic dynamics, and these differential dynamics may excite our various modalities in different ways, per the causal underpinnings of each. Many researchers addressing multimodal fusion methods tend to focus mainly on identifying new and better modalities and then integrating their inputs to improve performance, but they often tend to presume a symmetric relationship with the event in the world among the modalities. We believe that this unchecked symmetry presents a significant critique from both a statistical and first principles standpoint. This is critical, especially considering the high likelihood of underestimating the complexity of *building* a multimodal framework capable of modeling the true interdependencies among environment, sensor, and agents during complex real-world tasks. Within this light, we briefly evaluate the most common method of each general fusion approach, namely, *feature concatenation* (early fusion) and *weighted voting* (late fusion), because they compare with unimodal classifiers; all are applied to the challenge of detecting RWD events from human passenger behavior and physiology.

⁴ Ishmakhametov N, Naser MYM, Bhattacharya S. Assessing the rigor of machine learning in physiological signal processing applications. IEEE SoutheastCon; 2024. <https://doi.org/10.1109/SoutheastCon52093.2024.10500274>

D.2.1.1 Early Fusion by Method of Feature Concatenation

A general challenge for early fusion approaches is that they produce a dependency of the resulting feature vector on the richness of the individual modalities, which is necessary for successful measurement. However, this also means that the later fused results may be subject to influences not otherwise recoverable or removable. Different sensor modalities produce data that frequently are the net behavioral result of multiple, complex, underlying biophysical processes—some signals are known to be products of more complex underlying processes than others. For instance, EEGs detected at a single electrode are subject to the nonlinearly summed electromagnetic fields of literally tens to hundreds of thousands of individual neurons. Although eye movement is also produced by complex biological processes, it is nevertheless simpler; this is caused by the tensions on three sets of constrained muscle pairs that are themselves under the control of three dedicated cell groups in the brainstem, called nuclei, each evolved to specially manage particular aspects of optic function via movement. This means that although a sensor on each signal still effectively records a single variable and reports a single value each time instant (summed neural activity vs. gaze position), those variables may be validly decomposed into very different mixtures of fundamental data features that themselves have important relationships with aspects of the overall ongoing task. This is the first major concern with feature concatenation; that is, there is an inherent dependency of the resulting feature vector on the individual modalities and, specifically, the density and complexity of the biophysical sources contributing to the data they detect. When there are significant differences in the underlying features for each sensor input (e.g., number, complexity, and correlation of underlying features), one modality may dominate the other(s) in the final feature vector. This becomes more challenging in multiclass problems, such as those we face in the challenge of predicting braking, lane changing, and aggressive acceleration from the same base data.

As we view the configuration of the problem, we consider that different classes of events may be most appropriately correlated with specific modalities, or even particular groups of features from within modalities (consider spatial, temporal, and spectral domains). Capturing these more subtle associations offers a more nuanced approach to event classification that allows for variance in causal associations with underlying biological and physical phenomena. This issue is closely tied to the process of feature selection and processing; in short, determining the best way to prioritize feature contributions to yield accurate and statistically efficient classification. The second major consideration, given this more nuanced approach, is the allowance that different features have different meanings and operate on different scales. This also applies at the level of exact features; for instance, a gaze

position variability has a very different biological cause than heart rate variability. Therefore, in an analytic sense, it is necessary to establish an appropriate method for normalizing across features to eliminate and/or account for modality dominance as well as establish appropriate alignments.

Finally, to achieve a significant increase in classification accuracy, the modalities must contain features that exhibit complementary characteristics across the selected subset. This means that both *modality 1* and *modality 2*, for instance, *must* occur in relation to an event. In such cases, concatenating features is expected to provide substantial value, because it may enable the learning criterion to incorporate markers of that event from *all* modalities. On the contrary, in situations where the event can occur with *modality 1* and/or *modality 2*, combining the features is not expected to yield higher accuracies compared to the unimodal models. This is because the training assumes that both sets of features must exist for the event to occur, which is not always true in these cases. We have found that this distinction is often overlooked in the literature. The events' nature dictates, at least in part, these three considerations. Thus, an effective multimodal method should consider the type of events sought to be classified and their unique association with the sensed data. This is particularly crucial for multiclass problems like that represented in our RWD analysis. Therefore, we next evaluate a late fusion method that, as previously noted, methodologically attempts to model feature variances within modality with “fusion” occurring only after each sensor or feature type is processed into its contribution (“vote”) regarding the final output class.

D.2.1.2 Late Fusion by Method of Weighted Voting

Late fusion, in the classic sense, offers easier implementation compared to early fusion primarily because it does not require joint processing of different modalities. A weighted voting scheme is a late fusion method wherein each unimodal decision is quantitatively weighted according to certain rules or calculations and then, based on those weights, a final “fused” decision is made. For example, this approach was deployed in the proof-of-concept simulations that established the potential for ARL’s “Privileged Sensing Framework,” which likewise assumes that all agents who provide inputs contribute according to their own differential “privilege”-based weighting.⁵ However, the primary concern with this family of fusion approaches is that they treat each input separately and thus typically fail to leverage any complementary information that may be carried within each modality. This may be especially critical for biomedical applications of multimodal fusion, because every sensor records the same person who is responding to the same event at the same

⁵ Marathe et al. The privileged sensing framework: a principled approach to improved human–autonomy integration. *Theo Iss Ergon Sci.* 2018;19(3):283–320.

time; that is, there is certain information to be shared across all data modalities that reflect this one person. Furthermore, this voting approach brings no added value in cases of *modality 1 and modality 2*. Moreover, in the case of *modality 1 and/or modality 2*, late fusion methods may yield satisfactory performance improvements in very specific instances where 1) one modality is set as fundamental and 2) the other is supplementary to assist the decision of the superior modality. However, when multiple equally important modalities are present and when there is no a priori rationale for choosing a fundamental data modality, late fusion will offer no advantages because it then becomes a choice among modalities' inferring markers that may not be reliably present for each event that should be positively classified. Like early fusion, these scenarios are highly dependent on the event(s) under consideration, which is a more challenging task in multievent classification problems.

D.2.2 A Hybrid Solution: Sequential Event-based Multimodal Fusion

For both types of standard multimodal fusion approaches, there is a mutual dependence issue; the classification scheme relies on the unknown event being classified. To address this issue, in the following section we introduce a framework that we expect will alleviate the concerns.

D.2.2.1 Hybrid Fusion by Method of Sequential, Event-based Classes

The issue of mutual dependence between the classified event and the fusion mechanism, as discussed earlier, is unavoidable in classic multimodal schemes. However, this topic has received limited attention in the literature, possibly because it is perceived as an inherent challenge without a feasible solution, even though multiclass problems are arguably more prevalent than binary problems in general. This is particularly the case for physiological signal applications because of the interrelationships between them. Thus, we propose the hybrid scheme in Figure D-1; herein named the Sequential Event-based Multimodal Fusion (SEMF; developed in Neuro-Interaction Innovation Laboratory at KSU). This scheme offers a potential solution to effectively handle the simultaneous decoding of multiple decisions while efficiently consuming the limited data available for the task (e.g., from a single person). SEMF provides a framework for improved fusion in complex scenarios, provided access to complex multimodal data. In this scheme, the multiclass problem is approached using a sequence of single-event classifications based on subsets of data features from within the overall set. Each classification is treated as a decision level dedicated to the classification of a single event. This allows utilization of the most suitable set of multimodal data features per event, matched ideally based on fit between event and feature dynamics; in a sense, it pre-assigns a limited feature concatenation set for each event class.

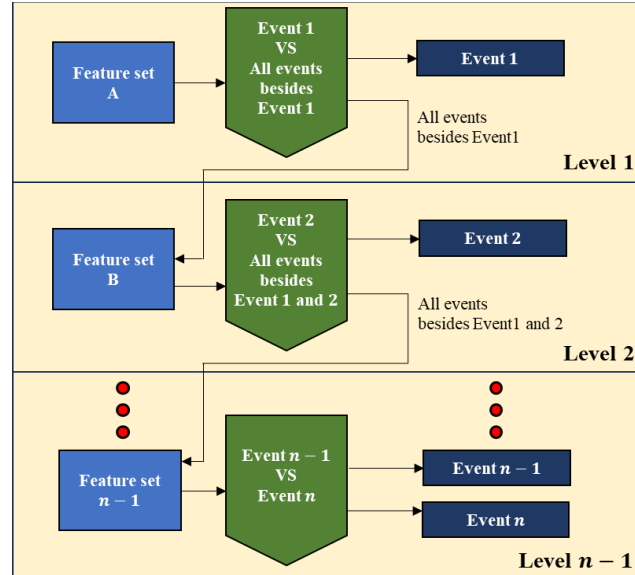


Figure D-1. Illustration of the SEMF approach.

This sequential approach allows for using the most appropriate processing steps to be applied for each event. SEMF relies on unique feature subsets at each decision level through an early fusion method of feature concatenation (discussed above). At each decision level, the feature vector comprises multiple modalities known to have verified data markers for the specific event of focus. Features are typically constituted by statistical combinations of mathematically extracted elements of the data series corresponding with one or more variables in the set.

The full classification of a data point, in the SEMF approach, is successfully complete when it is recognized as a target at any level or labeled as non-target. For instance, if the present data point belongs to the class “braking,” which is detected effectively by the feature subset assigned to decision level 2, then it will fail to be classified at decision level 1 and then pass on to decision level 2, where the process will be complete for that data point. That is, once the data point is recognized as belonging to a target class at any level, the classification decision is considered final for that point. Otherwise, the data point proceeds to each successive decision level. This process continues until all points are classified. The rationale behind the expected effectiveness of this approach is that each classifier within the SEMF hierarchy of decision levels is optimized to accurately classify a particular event; thus, the chances of misclassifying the data point are minimized while the feature data dedicated to that class will no longer be redundantly used at other decision levels.

We consider that SEMF does not belong to either early or late multimodal fusion families exclusively. Whereas the hierarchical set of decision levels is a late fusion method, in one sense, the deployment of feature processing and concatenation

within each decision level justifies our claim that this is a hybrid method, which is aimed to exploit the best elements of feature-based and decision-based approaches. The individual handling of events enables developers to manually adjust and prioritize decision-level performance toward specific events as needed, whether through qualitative constraint based on prior knowledge or through deployment of correlation methods and tuning or adjustment of statistical thresholds for aligning events to feature sets (i.e., such as forward and backward selection methods). Applying these approaches in a series of independent, single-event classifiers without the algorithmic control supplied within SEMF, however, risks improving classification quality on an individual event basis while compromising the quality of classification of the remaining events.

Returning one more time to our *modality 1 and/or modality 2* example, we note that SEMF benefits from essential features within both modalities as with the other approaches. However, with SEMF, the previously discussed troublesome case of *modality 1 and/or modality 2* is now managed through the applications of successive decision levels tailored to the relevant features for classifying the contextually situated event. In this way, a stringent criterion can be placed at the top decision level, where both modalities must all agree to classify a data point as the event. When this classification fails, more lenient criterion may be placed at successive levels where a less restrictive set may sufficiently classify the event, albeit perhaps with a lower confidence noted in classifier output. Unlike with other late fusion approaches, SEMF leverages the complementary information present in different modalities. In our observations, SEMF outperforms the approach of using a single set of features to classify one event at a time, without affecting the classifiers concerned with classifying other events in our initial applications to the RWD dataset; the advantage of this scheme may lie in the fact that downstream classifiers have fewer data points representing a smaller number of events, which may be expected to result in more successful classification process.

D.2.3 Preliminary Evidence: Sequential, Event-Driven vs. Multiclass Methods

To address our first concern, that is, whether multimodal would necessarily offer improvement over standard approaches, we randomly selected one drive from the full RWD dataset and identified the times at which each of the five event categories (braking, lane cross left, lane cross right, aggressive acceleration, and baseline “other” identifiable events that did not fall into any of the other four categories), which yielded a total of 257 events that might be related to momentary changes in any one of the features listed in Table D-3. Using early (concatenation) and late (majority voting) machine classification methods, we observed poor unimodal performance overall, with ECG providing the best event classifications (36.73%)

and eye-related features providing the worst (20.40%). To gain sensitivity through multimodal fusion, we noted slight improvement with overall classifications still unacceptably low; concatenation yielded a balanced accuracy of 38.05%; and our late fusion method, majority voting, performed at 45.32%. Indeed, given these initial results, we saw little value in proceeding to improve accuracy and optimize any of these approaches as a singular path forward. Rather, during this preliminary assessment, we also observed that some modalities were able to produce reasonable classification accuracies when applied in limited ways and focused on single events.

For an initial test case, using the same data discussed above, we performed “as-if” classifications where the multiclass problem was temporarily treated as independent, binary classification problems. As shown in Table D-4, the individual modality-event mappings produced some striking improvements on these more limited train-test runs, although still with data representing only a single participant. Test cases using the two multimodal approaches are included in the bottom two rows of Table D-4 for comparison. Two things we noted immediately were that 1) the unimodal and binary models almost always outperformed the multimodal models in at least one modality (matched in one case) and 2) with exception of the seat force sensors, which did not differentiate among unimodal events, each individual physiological modality seemed “best” suited for one of the events. EEG appeared to be a superior predictor of braking, the eye features were best for both acceleration and lane-changing behaviors, and ECG was a relatively close second for classifying aggressive acceleration.

Table D-4. Classifier accuracies (%) for uni- and multimodal methods applied to binary problems.

	Brake vs. baseline	Accelerate vs. baseline	Lane change vs. baseline
EEG	70	58	46
ECG	50	65	33
Eye	65	75	65
Seat	60	55	60
Early fusion	45	65	26
Late fusion	70	68	53

Emphasizing our previous argument, that the focus should be on event dynamics more than on identifying and then optimally processing multimodal inputs, we suggest that the data show the centrality of considering event dynamics as the major driver of classification performance. SEMF, as we have reviewed, deviates from more common, modality-centric practices encountered in the literature, because it prioritizes stability of classifier performance over established practices and traditions commonly adhered to within subdomains of biomedical science. Indeed,

appropriate fusion and multiclass solutions are necessary in appropriate contexts; moreover, we agree that more complete models of multimodal task dynamics are needed in general. However, given the original goal to successfully identify task-relevant events from human-sensed data, we argue that hybrid approaches provide usable class membership estimates with current state-of-the-art approaches

As a further justification to proceed developing SEMF, we performed one more set of classifications on the same passenger data discussed above, as well as two more randomly selected passengers. In this case, the same set of driving events were of focus, except now it explicitly treated all lane-changes as equivalent regardless of direction. For this classification task, across the three participants, there were 608 events ($n = 180, 257$, and 171 for passengers 1, 2, and 3, respectively) composed of 153 ($n = 53, 56$, and 44 for passengers 1, 2, and 3, respectively) braking, 243 ($n = 72, 97$, and 74 for passengers 1, 2, and 3, respectively) aggressive acceleration, 96 ($n = 19, 58$, and 19 for passengers 1, 2, and 3, respectively) lane changes, and 121 baseline events ($n = 36, 51$, and 34 for passengers 1, 2, and 3, respectively). Following removal of events associated with epochs containing corrupt physiological signals or other flawed data, the final event count was 598 (dropping seven epochs from passenger 1 and 4 from passenger 3). Now, the multi- and sequential-binary methods would be applied to an expanded set of potentially useful data features. In addition to the feature list from Table D-3, EEG activity in the “beta” component of the frequency spectrum was fractionated into three rather than two bands, ECG features were added to capture relative as well as absolute changes, and, finally, standard deviations variance measures were included alongside mean seat sensor output levels.

The results shown in Figure D-2 reinforce our conclusion that the multiclass formulation is outperformed by sequential application of selected feature subsets aimed at binary classification of single events such as braking and aggressive acceleration. In Figure D-2, approximate chance-level identification rates for braking and lane changes are indicated by the horizontal line just below 20% (light-blue dashed line), and the second line (horizontal yellow dashed line) is at 35% for the aggressive acceleration events. The degree to which the resulting classifier accuracies (represented with the vertical bars) exceed these thresholds provides qualitative support for our decision to more completely develop and pursue the SEMF approach for future applications involving complex, real-world datasets like that from the RWD study. Based on these initial evaluations, we concluded that a multiclass driving event prediction method will be insufficient (our classifications levels never exceeded 37% with any multiclass formulation). On the other hand, sequential, event-based feature selection and classification are supported by the

regular achievement of balanced accuracies as high as 66% by performing successive binary classifications with non-overlapping data subsets.

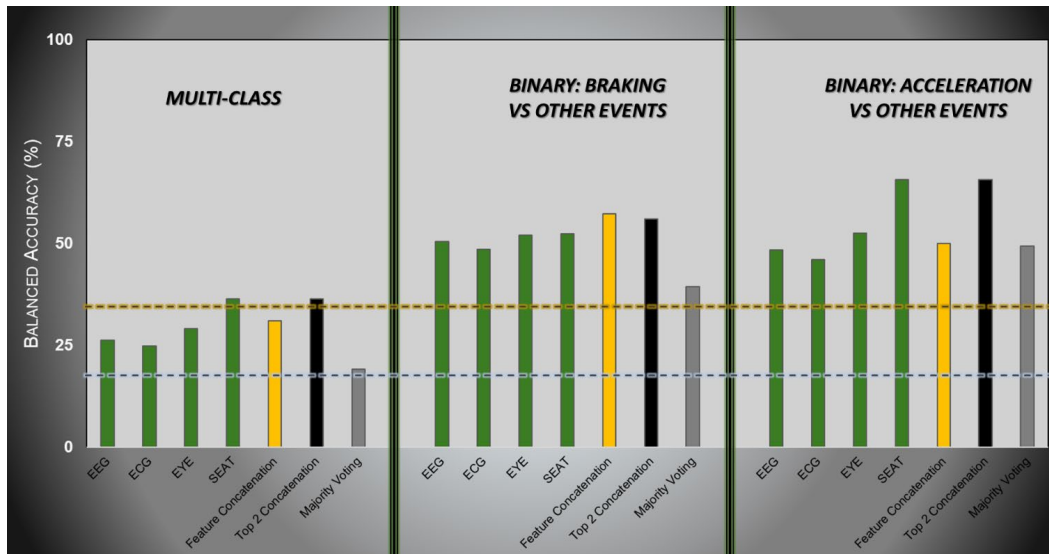


Figure D-2. Preliminary classification results for multiclass vs. sequential, binary methods.

The SEMF framework has yet to undergo thorough development, testing, and validation. However, the near doubling of accuracy on the same dataset that was seen here was striking and encouraging for our continued pursuit of this line of method development. Of course, much higher accuracies will be needed for any applications to real-world operations, especially those involving direct safety implications. Moreover, SEMF may also suffer disadvantages that can undermine performance if not handled properly. One disadvantage is the challenge of hidden or explicit dependencies between classifiers, such that weak initial classifiers would necessarily impact downstream performance in negative ways. A second major challenge is, thus far, we have not formally included any means of managing cases with very large numbers of data points, in general, as well as those datasets with imbalances in the event frequencies relative those of the actual world. Indeed, larger datasets risk a multiplicative increase in the number of floating-point operations required to achieve a final answer as to class membership and, therefore, an ultimate reduction of efficiency. To mitigate these issues, SEMF is envisioned to follow a hierarchical structure that always places the best and most complete classifiers at the top. In this way, the classifier may eliminate as many data points as possible while reducing the number of times each data point must be (re)classified. The ideal mix SEMF strategy, from this standpoint, is to select a range of classifiers with intentional combinations of higher- and lesser-performing classifiers used in the same decision hierarchy.

The above hierarchical strategy addresses our primary concern as well. This concern involves the well-known tradeoff between detecting events that happen (sensitivity) and accurately labeling events detected for what they are (specificity). In particular, SEMF was designed to enhance overall classifier sensitivity, which may have also increased odds of compromising specificity. By improving total number of detections, it becomes increasingly necessary to guard against the elevated likelihood of encountering false positives (e.g., detecting non-events as real). There are two options that we have considered to address this concern, and both fundamentally rest on pairing all classifier outputs with appropriately designed confidence metrics by Marathe et al.^{6,7} First, the SEMF algorithm can down-prioritize decision levels (model/feature set combinations) that perform at lower levels or with lower confidence, essentially placing them at a less consequential position in the SEMF classification queue without losing its incremental value added. Of course, this also inherently increases chances of more appropriate data decisions by the other classifiers before encountering the lesser model for that event or data point. Second, it is also possible to implement certainty checks through thresholds based on quantitatively determined confidence estimates on the decision outputs (at each layer). In essence, this places the SEMF on a convergent path with mentioned ARL efforts toward deploying advanced human–autonomy teams through evolving Army concepts such as the privileged sensing framework for human–autonomy teaming⁸ as well as opportunistic sensing for AI/ML applications.⁹

D.2.4 Next Steps: Applications to Human State Estimation in Risky Environments

As a major outcome of this project, future development of the SEMF approach will follow two courses. The first, mentioned in the main body of the attached report, will continue to be developed in a follow-on project also supported as a collaborative effort with ARL. In that project, we begin to design empirical platforms and paradigms that will yield data sets sufficient to explore multimodal

⁶ Marathe AR, Lance BJ, McDowell K, Nothwang WD, Metcalfe JS. Confidence metrics improve human–autonomy integration. Proceedings of the 2014 ACM/IEEE International Conference on Human–Robot Interaction; 2014 Mar. p. 240–241.

⁷ Marathe AR, McDaniel JR, Gordon SM, McDowell K. Confidence-based state estimation: a novel tool for test and evaluation of human-systems. Advances in Human Factors in Robots and Unmanned Systems: Proceedings of the AHFE 2016 International Conference on Human Factors in Robots and Unmanned Systems; 2016 July 27–31; Walt Disney World, FL. p. 291–303. Springer International Publishing.

⁸ Marathe AR, Metcalfe JS, Lance BJ, Lukos JR, Jangraw D, et al. The privileged sensing framework: a principled approach to improved human–autonomy integration. *Theo Iss Ergon Sci*. 2018;19(3):283–320.

⁹ Lance B, Larkin G, Touryan J, Rexwinkle J, Gutstein S. Minimizing data requirements for Soldier-interactive AI/ML applications through opportunistic sensing. *Proc. of SPIE*. 2020;11413:1141306–1.

fusion as a process that is computationally expensive and, therefore, only leveraged when necessary and, even in those moments, likely choosing to use only the amount and diversity of information needed to produce a behaviorally satisfying result. That is, we believe it would be very costly to design a system that continuously requires paying the cost of heavy cross-modality information fusion. Based on our own experiences developing multiclass methods for multimodal data, we imagine that there *are* types of information that are primary for particular tasks, and “fusion” only comes in when needed to resolve unexpected circumstances and outcomes. In the follow-up research, we develop paradigms and build technologies that allow precise control over task-related information, whether as various features embedded within a single modality (e.g., day vs. night processed visuals) or as combinations of fundamentally different modalities (e.g., video and directional audio). The empirical datasets will then be recorded on humans who will be given tasks that require periodic “arbitration” among various available pieces of information (some task-related and some not). In addition, the means by which humans dynamically choose new sources of information—when, for instance, the primary source has been made unreliable or unavailable (e.g., the lights are reduced to only 10% of their original luminance)—will be examined through close study of their behavior and the relevant physiology.

The second course of development focuses specifically on the core component needed to drive the SEMF method; that is, determining an algorithmic approach required to systematically deploy the hierarchical sequence of multimodal feature-based models in a manner that accurately identifies events of interest. A core assumption from the outset—again, deeply rooted in our experience with the RWD data set—is that the new method should allow for certain data, or their combinations, to be optimally deployed when a participant is expected to occupy a certain psychophysiological state, and, similarly, use other independent combinations when the person is in other states. To achieve this capability, we must better understand how to identify when a person is changing their state, noting the task or environmental triggers for the state change, and then we must determine how to systematically capture the dynamic bodily signals or, more specifically, the informative features within them that most powerfully signal each state. Indeed, this largely mathematical effort must ultimately converge with the first line of effort in the space of data. In particular, the testing of SEMF will require a new dataset consisting of human participants who have been skillfully taken on “controlled journeys” through a planned sequence of circumstances designed to induce different psychophysiological states. To enable consistent identification of this dynamic state, we leverage well-established paradigms based on fundamental state indicators, such as arousal and valence, which combine to yield more commonly known indicators such as “stress,” “frustration,” “anger,” “flow,” and so forth. With

a systematically labeled data, obtained through methods validated to shift states, analytic focus will assess modality or feature dynamics as a function of the valence-arousal state space, and machine classification will implement the results in our first formal test of the SEMF. Results are expected at the end of FY2025.

D.3 Code Catalog

This appendix outlines the software developed while producing evolving knowledge and technical products throughout the effort. Table D-5 delineates code groups as they are structured in our collaborative GitHub repository* and provides high-level descriptions of the contents of each file, with separate counts for total lines (LINES) and lines of code (LOC) contained within.

Table D-5. Code catalog.

Repository	File name	LOC/LINES	Operations	Owner
ECG	<i>CSV_UNIX_TO_MS.py</i>	45 /52	DO import TrialData SELECT passenger data DO assign ECG data DO set time arrays DO convert to milliseconds	M. Brown
	<i>Dyad1_Drive1_Breaking_vs_NoBreaking_ECG.py</i>	229 /260	DO convert to csv and SAVE BY format COMPRESSED	M. Brown
	<i>Dyad1_Drive1_HRV</i>	126 /141	DO import TrialData SELECT passenger, data DO assign ECG data SELECT braking, not braking DO filter and visualize WITH baseline removal DO plot Poincare	M. Brown
	<i>ECG_begin_aggressive_acceleration_session_1.py</i>	103 /124	DO import TrialData SELECT passenger, data DO assign ECG data DO HRV analysis ^a DO filter and visualize WITH baseline removal DO plot Poincare DO save csv with output	M. Brown
			DO <i>EventRelated Import.py</i> SELECT acceleration events DO R-R interval analysis WITH min-max normalization DO compute IBI, BPM	
EEG	<i>EventRelated Import.py</i> ^b	58 /81	DO import CSV file DO import TSV file FIND [relevant event tag] ** incorporated into MNE Code	M. Belcher

^a https://github.com/paulvangentcom/hearttrate_analysis_python/blob/master/examples/2_regular_ECG/Analysing_a_regular_ECG_signal.ipynb

* See repository at github.com/ARL-KSU

Table D-5. Code catalog (continued).

Repository	File name	LOC/LINES	Operations	Owner
EEG	<i>Signal Processing for EEG.py</i>	169 /209	DO Load EEG file SET data channel indices DO compute PSD BY Fast Fourier WITH Welch Method DO plot BY time, frequency ** incomplete CodeSnip	T. Dang
	<i>KarensMNECode.ipynb^c</i>	859 /1026 ^d	DO <i>EventRelated Import.py</i> DO import EEG RECORDING BY session SELECT data channels DO process each EVENT EXPLORE Welch method BY time AND epoch BY power AND epoch BY power ratio BY delta/beta, beta/alpha BY absolute power BY relative power BY multitaper method WITH 3 windows DO plot results	K. D'Souza T. Dang
	<i>MorgansMNEcode.ipynb</i>	648 /800	DO <i>EventRelated Import.py</i> DO convert event times to <i>ms</i> DO import EEG RECORDING BY session SELECT data channels DO process each EVENT SET epoch parameters DO gather data epochs SET Morlet parameters DO compute wavelets DO plot wavelet surfaces SAVE graphic outputs	M. Belcher I. Hwang
Head_	<i>OpenVINO_face-inference-</i>	35	DO face detection (CNN)	R. Kalidindi
Positioning	<i>results</i>	/52	WITH OpenVINO DO return head pose angles	M. Belcher

^b This code was inherited by KarensMNECode.ipynb, MorgansMNECode.ipynb, and ECG_begin_aggressive_acceleration_session_1.py

^c Leverages <https://mne.tools/stable/index.html>

^d These lines are in Jupyter Notebook (.ipynb), and thus they also contain developer-chosen screen text and graphic outputs. In another illustration of the conservation of intellectual effort in our team, this file also contains a significant number of code blocks (14 of 70) that preserve single-line errors that prevented the full and proper execution of some 503 remaining lines. On closer evaluation, those 14 errors directly affected 16 later lines that were dependent on their execution. In the end, errors (and warning flags also present) were points of learning, leading to evolved and more error-free software that later took this as its baseline, such as *MorgansMNECode.ipynb*.

Table D-5. Code catalog (continued).

Repository	File name	LOC/LINES	Operations	Owner
Head_ Positioning	<i>SVM</i>	108	DO Import libraries	R. Kalidindi
		/131	INCLUDE numpy, pandas matplotlib.pyplot sklearn.metrics DO read data FROM csv DO split(train, test) DO classify WITH kernel(linear) SHOW confusion, accuracy DO plot results	M. Belcher
Head_ Positioning (cont'd)	<i>correlate-with-driving-</i> <i>events</i>	263	DO import libraries	R. Kalidindi
		/334	INCLUDE tensorflow, pandas, numpy Imageio, os, csv, re matplotlib.pyplot imutils.paths DATA CAPTURE DO capture video frames AND crop, center, fit frame DO apply pre-trained network ^e BUILD feature extractor CONVERT strings to integers ASSEMBLE, EXECUTE SET sequence model (RNN) DO run model (train, test, eval) DO sequence prediction	M. Belcher
Seat Sensors	<i>final-code</i>	103	DO import libraries	P.K. Reddi
	<i>finalcode.ipynb</i>	/164	INCLUDE numpy, pandas, matplotlib, sklearn DO import data FROM Arduino DO compute (std. deviation) SET sensor column labels INCLUDE left, center, right DO transform WITH LabelEncoder DO split(train, test) CLASSIFY(SVM) WITH kernel(lear) DO cross-validation WITH k-fold (10)	Dendi

^e As in the code: The ['Keras Applications'](https://keras.io/api/applications/) module provides a number of state-of-the-art models pre-trained on the [ImageNet-1k dataset](http://image-net.org/). We use the [InceptionV3 model] (https://arxiv.org/abs/1512.00567) for this purpose.

Table D-5. Code catalog (continued).

Repository	File name	LOC/LINES	Operations	Owner
	<i>Configuration file.json</i>	69 /69	SET / MANAGE file paths, seed states training parameters	M. Naser
	<i>MMF code.py</i>	922 /1072	DO feature extraction see Table D-3 DO unimodal classification BY method of Concatenation BY method of Majority Voting DO Sequential Fusion (SEMF) SET training features DO feature selection BY statistical selection DO modality selection (sort) BY EEG, ECG, Eye, Seat DO split (train, test) CLASSIFY(SEMF) DO cross-validation BY k-fold (5) DO Exploratory Analysis	M. Naser
Multimodal fusion				

List of Symbols, Abbreviations, and Acronyms

2D	two-dimensional
ABDHS	ARL Big Data for Human Sensing
AI	artificial intelligence
APG	Aberdeen Proving Ground
ARL	Army Research Laboratory
Avg.	average
BCI	brain–computer interface
CAM	Cooperative Agreement Manager
CaN CTA	Cognition and Neuroergonomics Collaborative Technology Alliance
CAST	Center for Agent–Soldier Teaming
CNN	convolutional neural network
DEVCOM	U.S. Army Combat Capabilities Development Command
DoD	Department of Defense
ECG	electrocardiography
EEG	electroencephalography
EMG	electromyography
ERP	evoked response potential
fNIRS	functional near-infrared spectroscopy
FSR	force-sensing resistor
FY	fiscal year
HMI	human–machine interaction
IBS	Intentional Behavioral Synchrony
ID	identification
INFORMS	Information for Mixed Squads
KSU	Kennesaw State University

LINES	total lines
LLM	large language model
LOC	lines of code
Max	maximum
MEG	magnetoencephalography
ML	machine learning
MoBI	mobile brain–body imaging
M.Sc.	Master of Science
NIIL	Neuro-Interaction Innovation Laboratory
OpenVINO	Intel open toolkit for algorithm development
P300	positivity at 300 ms post-stimulus
PhD	Doctor of Philosophy
PI	Principal Investigator
RNN	Recurrent Neural Network, a structure used for propagating error in neural network training
RWD	real-world driving
S&T	science and technology
SEMF	Sequential Event-based Multimodal Fusion
SPSU	Southern Polytechnic State University
std. dev.	standard deviation
STEM	science, technology, engineering, and mathematics
STRONG	Strengthening Teamwork and Robust Operations in Novel Groups
SVM	support vector machine
UAV	unmanned aerial vehicle
UGV	unmanned ground vehicle
UMD	University of Maryland
WWII	World War II

1 DEFENSE TECHNICAL
(PDF) INFORMATION CTR
DTIC OCA

1 DEVCOM ARL
(PDF) FCDD RLB CI
TECH LIB