



ARL-TR-10306 • MAR 2026



Stages of Tactical Response Inferred from Kinematics and Eye-Tracking (STRIKE): Applications for Passive Targeting

by Leah R. Enders, Michael Nonte, Heather Roy, Louis Dankovitch, Russell Cohen Hoffing, Brandon Perelman, Thomas Rohaly, Thomas Dang, Alex Stauff, Jie Gao, and Jonathan Touryan

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.



Stages of Tactical Response Inferred from Kinematics and Eye-Tracking (STRIKE): Applications for Passive Targeting

Leah R. Enders, Heather Roy, Louis Dankovitch, Russell Cohen Hoffing, Brandon Perelman, and Jonathan Touryan
DEVCOM Army Research Laboratory

Michael Nonte, Thomas Rohaly, Thomas Dang, and Alex Stauff
DCS Corporation

Jie Gao
DEVCOM C5ISR Center

REPORT DOCUMENTATION PAGE

1. REPORT DATE		2. REPORT TYPE		3. DATES COVERED					
March 2026		Technical Report		<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;">START DATE</td> <td style="width: 50%;">END DATE</td> </tr> <tr> <td>5 June 2024</td> <td>9 February 2026</td> </tr> </table>		START DATE	END DATE	5 June 2024	9 February 2026
START DATE	END DATE								
5 June 2024	9 February 2026								
4. TITLE AND SUBTITLE									
Stages of Tactical Response Inferred from Kinematics and Eye-Tracking (STRIKE): Applications for Passive Targeting									
5a. CONTRACT NUMBER		5b. GRANT NUMBER		5c. PROGRAM ELEMENT NUMBER					
5d. PROJECT NUMBER		5e. TASK NUMBER		5f. WORK UNIT NUMBER					
6. AUTHOR(S)									
Leah R. Enders, Michael Nonte, Heather Roy, Louis Dankovitch, Russell Cohen Hoffing, Brandon Perelman, Thomas Rohaly, Thomas Dang, Alex Stauff, Jie Gao, and Jonathan Touryan									
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)				8. PERFORMING ORGANIZATION REPORT NUMBER					
DEVCOM Army Research Laboratory ATTN: FCDD-RLA-FC Aberdeen Proving Ground, MD 21005				ARL-TR-10306					
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			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: Leah R. Enders, 0000-0002-2205-2245; Thomas Rohaly, 0000-0002-1623-3015; Thomas Dang, 0009-0005-3468-6631; Jonathan Touryan, 0000-0001-7343-7869									
14. ABSTRACT									
ARL's Distributed Information for Enhanced Squad Lethality program aims to improve small unit lethality and situational awareness (SA) by transforming Soldier-system data into actionable information for humans and machines without imposing additional cognitive or physical burdens on Soldiers. This report details the Stages of Tactical Response Inferred from Kinematics and Eye-Tracking (STRIKE) study, which uses virtual reality and real-world experimentation to develop computational models that classify Soldier threat response states, quantify SA, and estimate weapon heading using kinematic and eye-tracking data. STRIKE employs inertial measurement units to capture head, torso, limb, and weapon movements during simulated tactical scenarios. As described in this interim report, preliminary results demonstrate the promising ability to classify distinct behaviors under nonthreatening and threatening conditions with key features such as elbow angles, knee angles, and the relative angle measures between the head, weapon, and rifle proving critical for model development. Additionally, STRIKE integrates data from multiple participants to explore passive pointing and covert targeting capabilities, offering potential applications for covert dismounted operations that require reduced Soldier signature and detectability. The study also introduces the Spatiotemporal Analysis and Inference Toolkit (STAIT), a standardized data-processing framework designed to synchronize and analyze heterogeneous data from mixed-reality devices and be useful for others working in this space, enabling robust algorithm development. Together, STRIKE and STAIT provide foundational tools for advancing human-machine integration and enhancing squad-level operational effectiveness.									
15. SUBJECT TERMS									
Humans in Complex Systems, opportunistic sensing, behavior detection, virtual reality, covert pointing, Soldier lethality									
16. SECURITY CLASSIFICATION OF:				17. LIMITATION OF ABSTRACT					
a. REPORT		b. ABSTRACT		18. NUMBER OF PAGES					
UNCLASSIFIED		UNCLASSIFIED		59					
c. THIS PAGE									
UNCLASSIFIED		UU							
19a. NAME OF RESPONSIBLE PERSON				19b. PHONE NUMBER (Include area code)					
Leah Enders				(520) 691-5504					

STANDARD FORM 298 (REV. 5/2020)

Prescribed by ANSI Std. Z39.18

Contents

List of Figures	v
List of Tables	vi
Executive Summary	vii
1. Introduction	1
1.1 Mission Description	1
1.2 Current State of the Art	2
1.3 Identification of Potential Stages of Tactical Response	2
1.4 Potential Use of Kinematics and Eye-Tracking Metrics to Construct Models of Threat Engagement and Tactical Response	4
1.5 Opportunistic Sensing with Inertial Measurement Units (IMUs) and Eye-Tracking Technology	5
1.6 Aims	7
2. Methods	8
2.1 Participants	8
2.2 STRIKE VR Visual Search Task	9
2.3 Equipment	10
2.4 Data Processing and Algorithm and Model Development	12
2.4.1 Overview	12
2.4.2 STAIT	13
2.4.3 IMU Feature Extraction	14
2.4.4 Segmentation of the IMU Feature Set and Application of Unsupervised Learning Methodology to Identify Preliminary Behaviors	16
2.4.5 Decision Tree Classifier Model to Detect Broad Weapon States	16
2.4.6 Building a Hierarchal IMU Feature Set Using IMU Feature Importance Analysis	17
2.4.7 Estimating Weapon Heading from the IMU Weapon Sensor	17

3. Results	19
3.1 Detecting Meaningful Behaviors in Our Unsupervised Learning Methodology	19
3.2 Successful Detection of Broad Weapon States Using a Decision Tree Classifier Using the Full IMU Feature Set	20
3.3 Identification of the SRSS	21
3.4 IMU-Estimated Weapon Heading Accuracy at Time-of-Shot and Triangulation on a Target	22
4. Discussion	26
4.1 Building the SRSS	27
4.2 Weapon Heading Estimates	28
4.3 Army and Future Applications	29
5. Conclusion	30
6. References	32
Appendix A. The Six Stages of Tactical Movement	38
Appendix B. Spatiotemporal Analysis and Inference Toolkit (STAIT) Overview	41
B.1 STAIT Toolbox Description and Usability	42
B.2 STAIT Toolbox Key Features	42
B.3 STAIT Toolbox Workflow	43
Appendix C. Full Inertial Measurement Unit Feature Set	44
List of Symbols, Abbreviations, and Acronyms	47
Distribution List	49

List of Figures

Figure 1.	Stages of tactical response and level of engagement based on weapon engagement and stability and alignment with the OODA loop.....	3
Figure 2.	STRIKE VR task environment examples: a) rooftop; b), first-person view with weapon pointed toward threat; c) friendly (left) and threat (right, weapon pointed) conditions in windows.	10
Figure 3.	a) Participants navigated the VR space in a large laboratory to allow free movement. b) Soldier with full experimental sensor setup including VR HMD, IMUs, and simulated weapon with embedded VR controllers.	11
Figure 4.	a) 15 IMUs were placed at head and joint locations. b) An additional IMU was affixed to the weapon c) to estimate weapon heading.....	11
Figure 5.	General overview of processing, with gold boxes showing outcomes demonstrated in this report.	13
Figure 6.	Data from single pilot subject shows estimated weapon heading from the IMUs compared with that from Unreal Engine controllers.....	18
Figure 7.	2D representation of all data points produced by UMAP showing at least two distinct clusters, indicating two distinct behavioral states within the no-threat game state for a selection of five features from the IMU feature set.	20
Figure 8.	Decision tree classifier found above-chance accuracy when compared with accuracy of shuffled weapon state labels obtained by bootstrapping (N = 100).	21
Figure 9.	Ten most important features for classification of weapon state (ranked) were the RAs of nine sensor pairs (Table 2) and the weapon's rotational velocity around the z-axis.	22
Figure 10.	Raw IMU weapon heading data from a single subject, demonstrating signal discrepancies when compared with the heading estimation from the Unreal Engine and VR controller.	23
Figure 11.	Correction of the IMU weapon IMU heading data improved alignment with the VR controller data from Unreal Engine for weapon course, pitch, and roll.	23
Figure 12.	Estimated weapon heading from two pilot participants' Unreal Engine data of a single threat attack were simultaneously cast into the environment.	25
Figure 13.	Triangulation error. Threat neutralized (red dash line) indicates when threat was hit (mitigated). Gaps in the data represent when one person's data was unavailable for a short amount of time.	26
Figure B-1.	STAIT pipeline.	43
Figure C-1.	Predicted importance weightings for the full set of features used in the STRIKE model. IMU = inertial measurement unit.	46

List of Tables

Table 1.	STRIKE VR scenario threat conditions and their descriptions.	10
Table 2.	Relative angles from nine sensor pairs (from 15 separate IMU sensors).	15

Executive Summary

To maintain overmatch in multi-domain operations, the U.S. Army must effectively integrate advanced technologies, including robotic and autonomous systems, at every echelon. The Soldier Lethality Cross-Functional Team (SL CFT) was established to meet this challenge by treating the Soldier and squad as an integrated combat system. Following this model, ARL initiated the Distributed Information for Enhanced Squad Lethality (DIESL) program. DIESL's core mission is to transform data from existing Soldier systems and subsystems into actionable intelligence. It achieves this through opportunistic sensing (obtaining operational data from tasks the operator is already doing¹) and leveraging a Soldier's natural physiological and behavioral data (e.g., movement) to infer their state and intent without adding to their cognitive or physical load in terms of size, weight, and power. This approach provides critical data for intelligent displays, command and control systems, and autonomous platforms, ultimately enhancing squad lethality and shared situational awareness.

A key application of this approach is the ability to infer a Soldier's position within the observe, orient, decide, act (OODA) loop based on their movements.² To achieve this, DIESL initiated the Stages of Tactical Response Inferred from Kinematics and Eye-Tracking (STRIKE) study. In this study, active-duty participants navigated a virtual environment and engaged threats under various conditions. By using a collection of inertial measurement units (IMUs) to track weapon, head, and limb movements, STRIKE aims to develop a computational model that accurately classifies a Soldier's response state within the detection, engagement, and assessment cycle to understand these movements and their relevance in operational contexts. The initial model informs the development of an optimized, field-ready sensor set (the STRIKE Reduced Sensor Set, or SRSS).

Conducting this research required a novel solution for managing complex data streams. Thus, DIESL developed the Spatiotemporal Analysis and Inference Toolkit (STAIT), a data format and software suite that synchronizes 3D time-series data from disparate physiological sensors and mixed-reality devices. As a shareable resource, STAIT provides a critical standard for this growing research area and was

¹ Lance BJ et al. Minimizing data requirements for soldier-interactive AI/ML applications through opportunistic sensing. In: Proc. SPIE 11413, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II; 2020 Apr 23 [online only]. SPIE; 2020. p. 1141306. <https://doi.org/10.1117/12.2564515>

² Howe PR. Leadership and training for the fight: a few thoughts on leadership and training from a former special operations soldier. AuthorHouse; 2005.

used to process all data in the present analysis. This report serves as an introduction to the STAIT toolbox, with STRIKE data serving as the first use case.

This interim report presents preliminary STRIKE study results ($N = 16$) demonstrating early success. Using clustering methodologies on kinematic movement data, we can distinguish meaningful behaviors under different conditions. In nonthreatening scenarios, we can classify distinct "search" and "movement" phases. In threatening scenarios, we can classify behaviors around target engagement (weapon lowered, weapon raising, weapon raised, weapon lowering) with greater than 61% accuracy, primarily using relative angles from the weapon's IMU. These behaviors directly correspond to the observe and orient phases of the OODA loop during search, and the decide and act phases during engagement, linking observable motion to the cognitive decision-making process. Lastly, we identified a preliminary SRSS in our preliminary model that includes the torso, head, and weapon IMU. Future work with the completed STRIKE dataset will focus on refinement of these models to further improve accuracy and real-time capabilities, finalize the SRSS, and include eye-tracking data to assess potential model performance enhancements.

Looking forward, this methodology offers proof of concept for passive and covert targeting. Although STRIKE data was collected from single participants (serial collection), we can integrate data from multiple participants into a shared threat environment for parallel analysis. This provides a unique opportunity to address specific challenges around passive pointing and covert targeting. The desire for passive/covert pointing technologies is becoming increasingly important in military operations for detecting targets and areas of interest, tracking, and localization without electromagnetic emissions, offering significant stealth advantages and reducing Soldier signature and increasing survivability. While many current efforts focus on equipping the individual Soldier with this capability (e.g., a passive pointer/marker), our analysis demonstrates a multi-Soldier approach. By showing how pose, position, and heading estimations from multiple weapons can be fused to pinpoint targets as Soldiers naturally converge on them, this work highlights a path toward reducing Soldiers' electromagnetic signature and increasing survivability on the battlefield.

1. Introduction

1.1 Mission Description

The Army seeks to discover novel mechanisms to provide context and insight into how individuals and teams understand complex real-world environments so intelligent systems can effectively communicate and adapt, a foundational component of effective human-machine integration (HMI). Currently, human-system (i.e., human and intelligent technologies such as unmanned aerial/ground vehicles) performance is limited as intelligent systems, where they exist, lack the capability to robustly predict human behavior, actions, goals, and intentions, limiting their capability to integrate with humans (Chen and Barnes 2014; National Academies of Sciences, Engineering, and Medicine 2022; Lou et al. 2025). Humans are superior to autonomous systems in their ability to communicate, intentionally or unintentionally, to one another via a myriad of physical and verbal indicators (Cartmill et al. 2012; Siposova et al. 2018). Furthermore, physical movement itself provides a nonverbal form of communication and, when observed or quantified, provides a dense source of information for understanding the dynamic mission context. Here, we refer to the process of unobtrusively obtaining and exploiting this information as “opportunistic sensing.”

Specifically, opportunistic sensing refers to the process of leveraging the human as a sensor through the nonintrusive collection and analysis of Soldiers’ physiological and behavioral data while they naturally perform mission-relevant tasks. Opportunistically sensed data can be used to infer mission context, train predictive algorithms, and enhance unit situational awareness (SA) and effectiveness. Through opportunistic sensing of physical movement, we can create novel methodologies for effective and efficient real-time, multimodal communication between Soldiers and intelligent systems. If postural information from a Soldier can be sensed and interpreted in real time, this information can yield vital battlefield information such as threat(s) presence, location(s), and intensity. Likewise, combining information with varying perspectives and sensing capabilities from other teammates and intelligent systems distributed across the battlefield could provide a more holistic view of the environment and critical battlefield information to commanders and autonomous systems.

This report uses decades of research on the biomechanics of motion and posture, as well as visual attention and search in natural environments, in an effort to determine if the stages of tactical response can be identified through an opportunistic sensing framework. We combine state-of-the-art virtual experimentation and biomechanical analysis to develop robust inferential algorithms that generalize

across participants and use the ground truth provided by virtual experimentation to explore tradeoffs between density of information and accuracy of classification or estimation. This work uses the growing technological capabilities of multimodal sensors, spatial computing systems, and inertial navigation algorithms, ultimately seeking to identify a minimal set of fieldable sensors that can classify Soldier state and weapon heading in operational environments. While the utility of this data is manifold, including hands-free activation of scope imaging technologies, the present study focuses on applications for passive pointing or passive threat identification and localization across the squad.

1.2 Current State of the Art

ML techniques have rapidly advanced in their ability to interpret human state during movement, enabling this state data to be integrated into the development and training of intelligent systems (Halilaj et al. 2018; Horst et al. 2019; Zago et al. 2021). These systems are now able to recognize general movement patterns, postures, and even predict movement intent in controlled environments (Zaroug et al. 2020; Kolaghassi et al. 2023; Hollinger et al. 2024). However, such approaches remain limited in their applicability to support integration into military Soldier–intelligent systems due to current systems’ lack of ability to accurately interpret and predict Soldier state during tactical movements, as these movements are dynamic and complex. This problem limits our real-time capabilities of integrating intelligent agent systems into Soldier teams and begins with a lack of tactical movement data available to train ML models. To this end, we developed our study and testing platform, Stages of Tactical Response Inferred from Kinematics and Eye-Tracking (STRIKE), to capture natural Soldier movement and response during a tactical scenario.

1.3 Identification of Potential Stages of Tactical Response

The tactical readiness cycle describes the stages of movement that a Soldier goes through when detecting, mitigating (shooting and removing threats), and confirming the mitigation of a threat. Soldier movement during a tactical engagement is a series of complex actions that can be understood through the lens of the observe, orient, decide, act (OODA) loop—a foundational model for understanding the phases of rapid decision-making (Howe 2005). The tactical readiness cycle can be classified into six basic stages (Figure 1) that map directly to the OODA framework: low threat, threat alert, threat detected, threat identified, threat engaged, and confirmation (HDA 2016a). The identification of these stages is critical for developing predictive models and algorithms. See Appendix A for

detailed descriptions of each stage of tactical response and the specific physical and weapon movements associated with each stage.

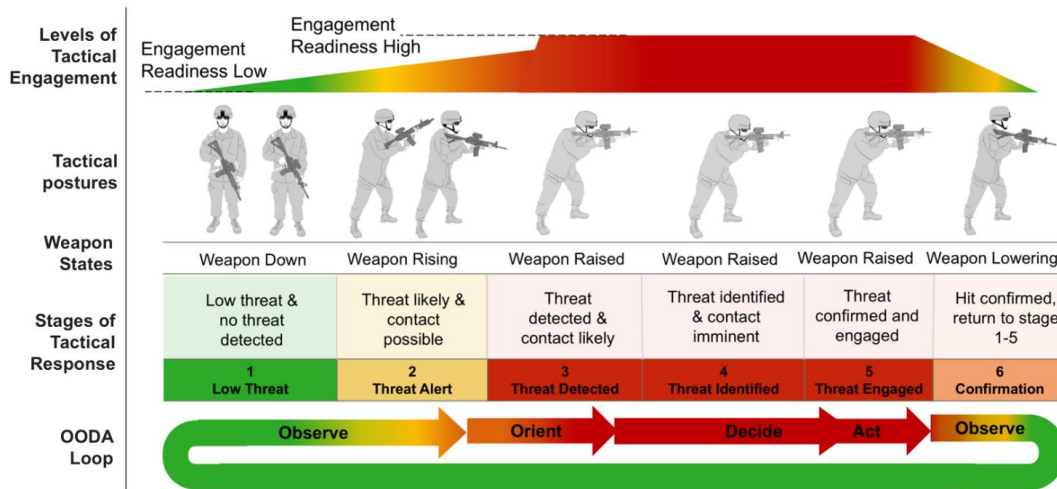


Figure 1. Stages of tactical response and level of engagement based on weapon engagement and stability (HDA 2016a) and alignment with the OODA loop (Howe 2005).

The observe phase encompasses both the initial search for threats and the moment a threat is identified, mapping onto the “low threat” and “threat alert” stages. Once a threat is observed, the Soldier transitions to the orient phase, which corresponds to the “threat detected” phase within the STRIKE model as they process the threat’s location and context. The “threat identified” and “threat engaged” stages represent the tightly coupled decide and act phases of the OODA loop. Finally, the “confirmation” stage serves as a new observation, providing critical feedback that initiates the next iteration of the loop. Although cyclical in nature, progression is not always linear; a Soldier may move backward and forward between stages as their readiness to fire increases or dissipates.

STRIKE models this process for a midrange threat. Thus, this tactical response cycle is not inclusive of all threat-contact scenarios; the work included in STRIKE models a response close to when a Soldier would be first engaging a midrange-type threat. A midrange threat, roughly between 30 and 100 m away, and is too far away for close combat procedures, yet too close for downrange threat-engagement tactics (e.g., firing from a prone supported position). Investigating a midrange target scenario is the initial focus of STRIKE modeling efforts and algorithm development, and these stages of tactical response are not inclusive of all postures that may occur.

1.4 Potential Use of Kinematics and Eye-Tracking Metrics to Construct Models of Threat Engagement and Tactical Response

Kinematic movement, gross and fine, can be used as postural indicators of stages of Soldier response and may be sensed opportunistically in the field. Head movement, not only a marker of engaging in active search activity, but also head turning (e.g., rotation about pitch and roll axes) are primary indicators of impending whole-body directional heading changes (Imai et al. 2001). Head turning and shifts in eye movement (see paragraphs directly below) act together as the feed-forward navigational control system that precedes lower-body movement during walking and turning. These can be early indicators of an impending heading change (Grasso et al. 1998; Imai et al. 2001; Hollands et al. 2002), useful in objectively classifying tactical response state changes of threat alert, threat detection, and threat identification. Grasso et al. (1998) found this anticipatory movement of the head and gaze shifts occurred approximately 1 s prior to directional shifts during walking.

Lower-body indicators are also important indicators of shifts in tactical response. There are notable differences in spatiotemporal parameters of gait and lower body kinematics between normal walking and dynamic shooting (i.e., walking while holding the weapon aimed) (Talarico and Neugebauer 2017). Dynamic shooting postures, such as rolling the feet while walking, limiting movement of the hips, and rotating at the waist (i.e., creating a “turret-like” movement of the upper body), work in conjunction to keep the raised rifle steady during forward movement in active tactical engagement (HDA 2016a). Kinematic measurements of torso–rifle linkage (i.e., ratio of the rifle and torso angular velocities, capturing joint movement) and comparisons of rifle and torso stability (i.e., maintaining low upper-body and limb variability in the linear acceleration) capture this “turret-like” movement (O’Donovan et al. 2024). Additionally, arm movement trajectories are dependent on displacements occurring in the center of mass (Pozzo et al. 2002) and capturing these differences of when Soldiers engage or move between these types of postures could be useful classification features of tactical response. Furthermore, research has shown physical domain features including upper- and lower-body joint strength, range of motion, and flexibility are important physical components for explaining variance in Soldiers’ operational performance (Giles et al. 2023). Given this work, head and upper- and lower-body kinematic measurements could inherently be a rich resource for capturing changes in Soldier readiness during the tactical response cycle.

In addition to kinematics, gaze dynamics can also yield insight into classifying stages of Soldier response towards threat(s) in addition to proximal measures of SA. As mentioned earlier, gaze shifts accompany head movement as early indicators of whole-body heading changes where fixations shift to focus in future directions (Grasso et al. 1998; Imai et al. 2001; Hollands et al. 2002). Eye-tracking in general is a common objective assessment for predicting SA (Zhang et al. 2023), giving insight into the state of one's knowledge pertaining to the state of a dynamic environment (Endsley 2024). SA involves the timely perception of key elements in the environment, comprehension of their meaning, and the ability to project implications for the near future (Endsley 1988, 1995). Previous work in our lab has shown eye-gaze with an immersive desktop visual search task provides critical information regarding attention (Enders et al. 2021) and has identified Soldier-specific visual search patterns including increased peak velocities accompanied by greater average saccade magnitudes (Enders et al. 2024), a behavior linked with expert visual search strategies (Castner et al. 2022). This work showed that when searching for targets in desktop virtual environments, participants significantly increased fixations and dwell time on targets compared with distractor items in the environment, indicating they prioritized target information in the environment (Enders et al. 2021). When cognitive load was increased with an additional auditory task, fixations and dwell time were reduced on objects in general (Enders et al. 2021). This shows how increased attentional demand can impact SA and scanning behaviors.

Eye-tracking provides indicators for visual recognition, cognitive decision-making, and attention, and could be preemptive of postural adjustments in terms of threat recognition and detection. These measures are of exceptional value when assessing expertise in tasks reliant on hand-eye coordination (e.g., weapon firing and medical procedures). Inclusion of eye-tracking measures in models trained to predict expertise in the performance of medical procedures significantly improved accuracy by 5.3% and 13.2% for novices and experts, respectfully, compared with models trained on kinematic data alone (Ahmidi et al. 2012; Fox and Faulkner-Jones 2017). Incorporating both whole-body kinematic movements and eye-tracking will increase the robustness of models of tactical response and engagement compared with those offered by eye-tracking or joint kinematics recordings alone.

1.5 Opportunistic Sensing with Inertial Measurement Units (IMUs) and Eye-Tracking Technology

Building threat response models from kinematics and eye-tracking data provides a potential pathway for tactical applications. Current and future Soldier systems will likely include sensors that capture this data or some proxy of reasonable fidelity

(e.g., head movement). Although historically restrictive and cumbersome to capture, advances in wireless motion capture technology (e.g., IMUs) have broadened the experimental scope through which we can capture kinematics during tasks such as dynamic weapon firing. IMUs combine accelerometers, gyroscopes, and sometimes magnetometers to continuously measure the acceleration, rotational velocity, and orientation of an object (or limb). When placed on neighboring body segments (e.g., upper arm and forearm), joint angles (e.g., elbow flexion and extension) can also be computed. Because this system is wireless and not dependent on external cameras such as the more restrictive, traditional, optical marker-based motion capture systems (e.g., Vicon), IMUs offer a more mobile measurement system for tracking kinematics, especially with in-field applications. McGinnis et al. (2014) found that torso pitch angles, rifle elevation, and azimuthal angles were comparable between traditional marker-based motion capture and IMUs with small angular differences of 0.44° – 1.26° , on average, during an up–down maneuver. Mavor et al. (2020) found an acceptable average correlation coefficient (0.81 ± 0.14) between measurements captured by IMUs and an optical motion capture system in tasks involving a broad range of military maneuvers with larger errors in angular estimates for upper-body joint angles.

Errors in joint angle estimations tend to increase with IMUs over time (even in short collection durations) due, in part, to estimations of the position or joint angle relying on multiple integrations of the acceleration-based data, which leads to the possibility of error inflation. However, such issues estimating joint angles can be minimized with correction models, some of which are done in the IMU device-specific preprocessing package for output of the joint angle data (e.g., Noraxon MyoMotion). Alternatively, one could assess postural changes directly via the accelerometry data. O'Donovan et al. (2024) used IMU acceleration and angular velocity data directly, comparing changes in segment linkages (e.g., torso–rifle linkage) and stability to determine postural differences between experts and novices during a dynamic firing task. Additionally, IMU time-event signals can readily be aligned with other physiological signals (e.g., eye-tracking) and virtual reality (VR) system events, offering an opportunity to capture kinematics dynamically in multimodal studies.

Gaze information can provide rich insight into cognitive processes in both laboratory (König et al. 2016) and military contexts (Enders et al. 2024). Typically, gaze direction is captured through imagery of the eye and estimation of pupil orientation via some form of IR illumination. Measures of gaze direction are then divided into features such as saccades, fixations, and blink, each with a corresponding series of metrics. These features can then be used to develop models of task performance or perceptual constructs such as SA. Despite advances in

camera-based mobile eye tracking, capturing relative or absolute gaze angle across varied real-world environments remains challenging due to factors like variable illumination or device slippage (Hessels et al. 2020; Ghiani et al. 2023). In contrast, head-tracking sensors, such as accelerometers or IMUs, offer a more robust—while less precise—means to estimate gaze direction, making them an attractive alternative for capturing data in challenging environments.

Our recent findings underscore the potential of head-tracking data to quantify task-induced changes in behavior, even when eye-tracking data are unavailable (Ries et al. 2025). Furthermore, the simplicity and durability of head-tracking sensors make them a practical solution for applications requiring lightweight, unobtrusive, and cost-effective systems. Prior research has demonstrated the utility of head kinematics for inferring search behavior (Hu et al. 2023; Salehi et al. 2024). Hu et al. (2023) showed head movement velocity and acceleration could reliably distinguish between different types of visual search such as free viewing, target search, and tracking. Similarly, Aivar et al. (2024), demonstrated body movement velocity increases with repeated search in the same context, reflecting an optimization of movement strategies for faster, more efficient search execution. Importantly, head-mounted display (HMD) systems, both commercial and military, typically include head-tracking technology. This is critical for the stability of the visual display in the context of head movement. In particular, the Army’s Soldier Borne Mission Command (formerly known as the Integrated Visual Augmentation System) has head-tracking capabilities, with iterations containing active eye-tracking capabilities and future iterations that may contain passive capabilities.

1.6 Aims

In this study, participants with tactical weapons performed a surveillance and threat mitigation VR rooftop task while kinematic IMU sensors were affixed to the head, torso, limbs, and rifle. Threats included friendlies, informed threats (instruction of threat impending attack and general location of threat), and uninformed threats (no indication of threat presence or location). This interim report (with 16 subjects) discusses preliminary work related to development of models that will inform the development of an optimized, field-ready STRIKE Reduced Sensor Set (SRSS). Initial models in this interim report only included IMU kinematic data; however, additional features from other sensing modalities (e.g., eye tracking collected from VR) will be assessed in the future to improve model performance.

The overall aims of STRIKE are twofold:

1. Develop a model using kinematic data collected from IMU sensors that accurately classifies weapon state (i.e., position and motion) related to the stages of tactical response; and
2. Determine an SRSS that classifies stages within the threat response cycle by investigating the sensitivity of each sensor and feature to overall classification accuracy.

In addition to these aims, we explored moving beyond the individual by combining STRIKE data from multiple individuals to triangulate areas of interest to inform the development of passive or covert pointing technologies. This report presents our initial capabilities.

The STRIKE dataset serves as the foundation for the development of the Spatiotemporal Analysis and Inference Toolkit (STAIT). STAIT was developed as a toolkit to transform data into a homogenous format and provide capabilities to synchronize 3D time-series data from disparate physiological sensors and mixed-reality devices (see Section 2.4.2). Both STRIKE and STAIT were developed in conjunction to serve as foundational work for future studies and algorithm development.

2. Methods

2.1 Participants

STRIKE is recruiting up to 40 active-duty Soldier participants from Aberdeen Proving Ground (APG); Adelphi Laboratory Center; U.S. Army Combat Capabilities Development Command (DEVCOM) C5ISR; and the Delta Company, First Battalion, Third Infantry Regiment. These Soldiers must have tactical weapons experience and approval from their command to participate. Tactical weapons experience is defined here as those having received tactical threat training or those that have had sufficient experience in tactical movement with a rifle or simulated rifle. Specifically, training in how to maneuver strategically with a weapon, in a coordinated manner, to gain tactical advantage against a threat by utilizing any (and all) cover, concealment, and positioning advantages, while minimizing exposure to enemy fire. Proper rifle handling during dynamic shooting is a skill developed through experience and training, and a model predicting Soldier responses should be trained on data from stereotypical postural movements and adjustments associated with those trained in tactical response. As of this writing, 16 active-duty Soldiers completed STRIKE, and there are 13 total datasets (one was lost to poor data quality and two did not qualify based on inclusion and exclusion

criteria). All participants were volunteers who gave informed consent to take part in the study. The protocol for the study (DEVCOM Army Research Laboratory [ARL] 25-028) was approved by the ARL Institutional Review Board.

This study was based on a pilot study (ARL 24-044) that followed the same study activities and setup as the STRIKE full study. However, the pilot study volunteers (N = 11) were civilians who did not necessarily have tactical weapons experience, and the data was used only to establish construct validity and for preliminary STAIT algorithm development.

2.2 STRIKE VR Visual Search Task

The STRIKE visual search task required participants to stand and physically navigate (ambulate) in a virtual rooftop urban environment with buildings surrounding their rooftop location (Figure 2). This virtual environment was developed in the Unreal Engine (a 3D game engine from Epic Games, Inc.). Participants were tasked with patrolling the rooftop to provide surveillance and reconnaissance. As shown in Table 1, they completed three experimental conditions: 1) no threats present, 2) threats present with audio alert (threat condition 1, informed threat), and 3) threats present with no verbal instruction (threat condition 2, uninformed threat). Threat conditions were designed to provide a range of threat scenarios in a controlled experimental paradigm but also include the unpredictability and variation one may encounter during such a task in the field. The paradigm included 24 trials total, with each trial including either a singular (50%) or mix (50%) of conditions. The reasons for this are twofold: 1) singular condition trials (only one nonthreat or threat type) are needed for training the classification algorithms, and 2) maximizing variance and avoiding preemptive or stereotyped responses. Trial types were presented (seemingly) randomly to participants. Each trial was approximately 1–2 min in duration. Participants' participation in the study took approximately 2.5 h.

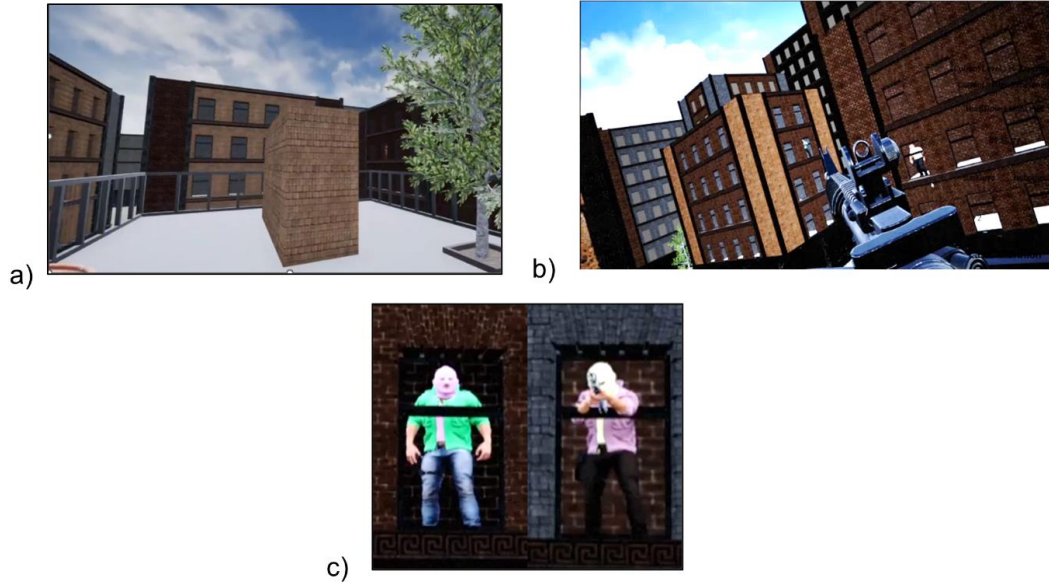


Figure 2. STRIKE VR task environment examples: a) rooftop; b), first-person view with weapon pointed toward threat; c) friendly (left) and threat (right, weapon pointed) conditions in windows.

Table 1. STRIKE VR scenario threat conditions and their descriptions.

Condition	Type	Description
No threat condition	Friendly	No threats occur. Only friendlies with no weapons raised appear in building windows. Participants are to remain in the low threat/walking (patrol operating) state while physically patrolling rooftop area.
Threat condition 1	Informed threat	Participants provided with verbal alerts indicating threat and approximate threat location (e.g., “Threat in the northwest”). Threat appears with weapon pointed.
Threat condition 2	Uninformed threat	Participants do not receive any verbal alerts to threat presence. Threat is known either through visual search of environment or due to sounds of simulated gunfire directed towards participant. Gunfire sound is not binaurally localized, limiting spatial information about threat and encouraging search behavior. Threat appears with weapon pointed.

2.3 Equipment

Data collection was conducted at an indoor, climate-controlled APG laboratory (Figure 3a). The large space provided a clear, safe area to accommodate free physical movement of participants while they navigated the VR environment. The space also provided an unobstructed, complete view for researchers to monitor participants while they were engaged with the virtual environment. A Quest Pro headset (Meta) presented the VR scenario environment and audio and collected eye-tracking metrics (Figures 3b and 4a). For immersive simulations, it is critical that the shooting interface physically resembles the weapon being evaluated in both

form and weight so the same aiming and firing techniques are used (Kehring 2012). Thus, an airsoft rifle was used to mimic a military-grade rifle (e.g., M4) (Figure 4b). The controller was mounted to the weapon stock at the approximate location for natural trigger pull. Additionally, a weight was added to the weapon to approximate the weight of an actual military-grade weapon. These modifications help facilitate natural kinematic movements associated with tactical response.

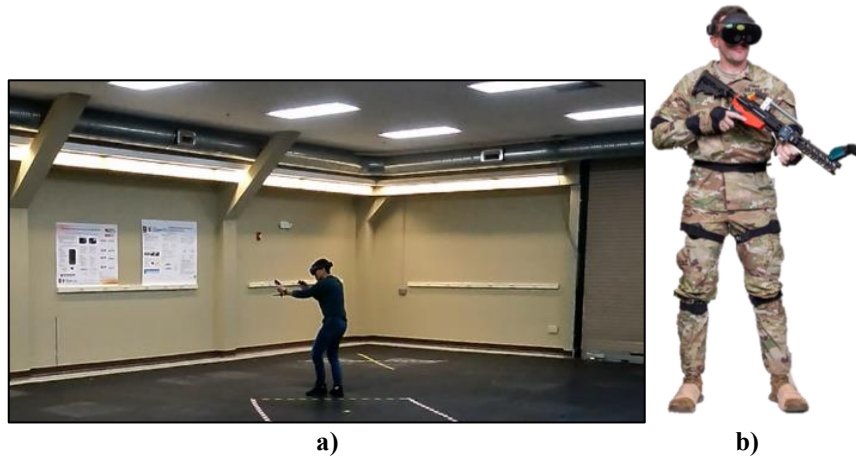


Figure 3. a) Participants navigated the VR space in a large laboratory to allow free movement. b) Soldier with full experimental sensor setup including VR HMD, IMUs, and simulated weapon with embedded VR controllers.

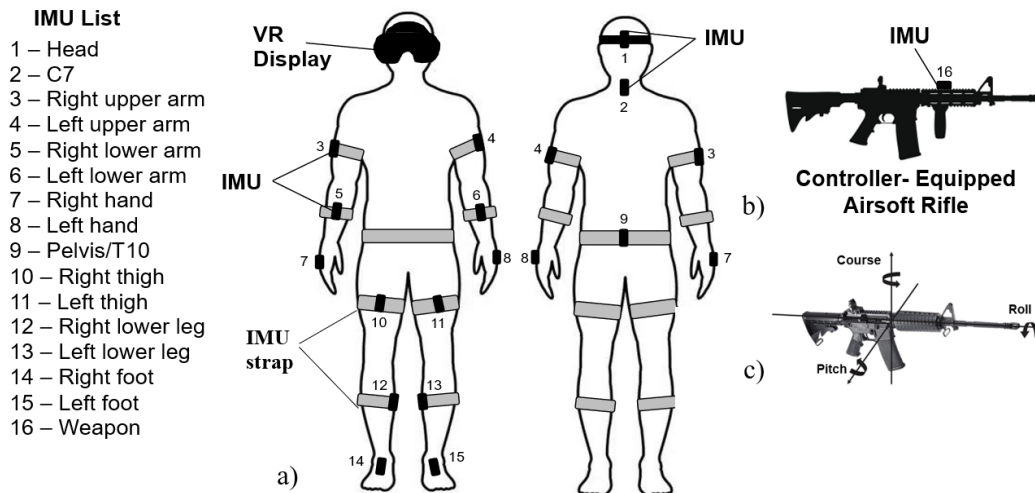


Figure 4. a) 15 IMUs were placed at head and joint locations. b) An additional IMU was affixed to the weapon c) to estimate weapon heading.

To capture participant movements, a 3D Ultium Motion Capture System comprising 15 Noraxon IMUs placed on the head, torso, and limbs provided anatomical joint angles, orientation angles, and limb linear acceleration (Figures 3b and 4a). An additional IMU was placed on the rifle to provide weapon position and orientation (Figure 4b). Each IMU possessed 9 degrees of freedom (9DoF; three-axis accelerometer, three-axis gyroscope, and three-axis magnetometer) recording capabilities.

During recording, a synchronization pulse was generated by an Arduino microcontroller that simultaneously sent digital markers coinciding with pulse edges to the data collection PC. These data were captured using the Lab Streaming Layer protocol and LabRecorder application (University of California, San Diego). This enabled synchronizing the IMU system with all other experimental data. All VR data (headset, controllers, game events, etc.) were acquired through the Unreal Engine and SteamVR middleware (Valve Corporation) using UETack, a custom plugin previously created by ARL that captures and logs all virtual objects, frame renders, game events, and device movements (Touryan et al. 2023). Participants' movement in the laboratory space was also visually recorded with a standard GoPro Hero9 camera. This video will be used to help classify physical movements of the participants and to train and confirm model estimation of tactical response stages and postures.

2.4 Data Processing and Algorithm and Model Development

2.4.1 Overview

Data from multiple sources was first processed and then used to develop algorithms and models for classifying weapon state (Figure 5). To start, data was spatially and time synchronized across the VR headset and controllers, IMU sensors, and Unreal Engine. Using the time-synchronized game event tags, the IMU feature set data was sectioned to include only data during a time when no threat was present. Unsupervised learning methodology was applied to this sectioned data to demonstrate the capability of detect distinct behaviors with this IMU feature set and to begin to understand underlying behaviors that occur during this “surveillance” nonthreat stage. A decision tree classifier using the full IMU feature set (including threat and nonthreat times) determined if, using these sensors holistically, it was possible to determine broader stages of tactical readiness (weapon up or down). Then, a MATLAB Feature Importance methodology determined which sensors were most important in distinguishing these broader stages of tactical readiness. Finally, some preliminary proof-of-concept results on

how STRIKE sensors could be used for approximating outcomes such as weapon heading are showcased.

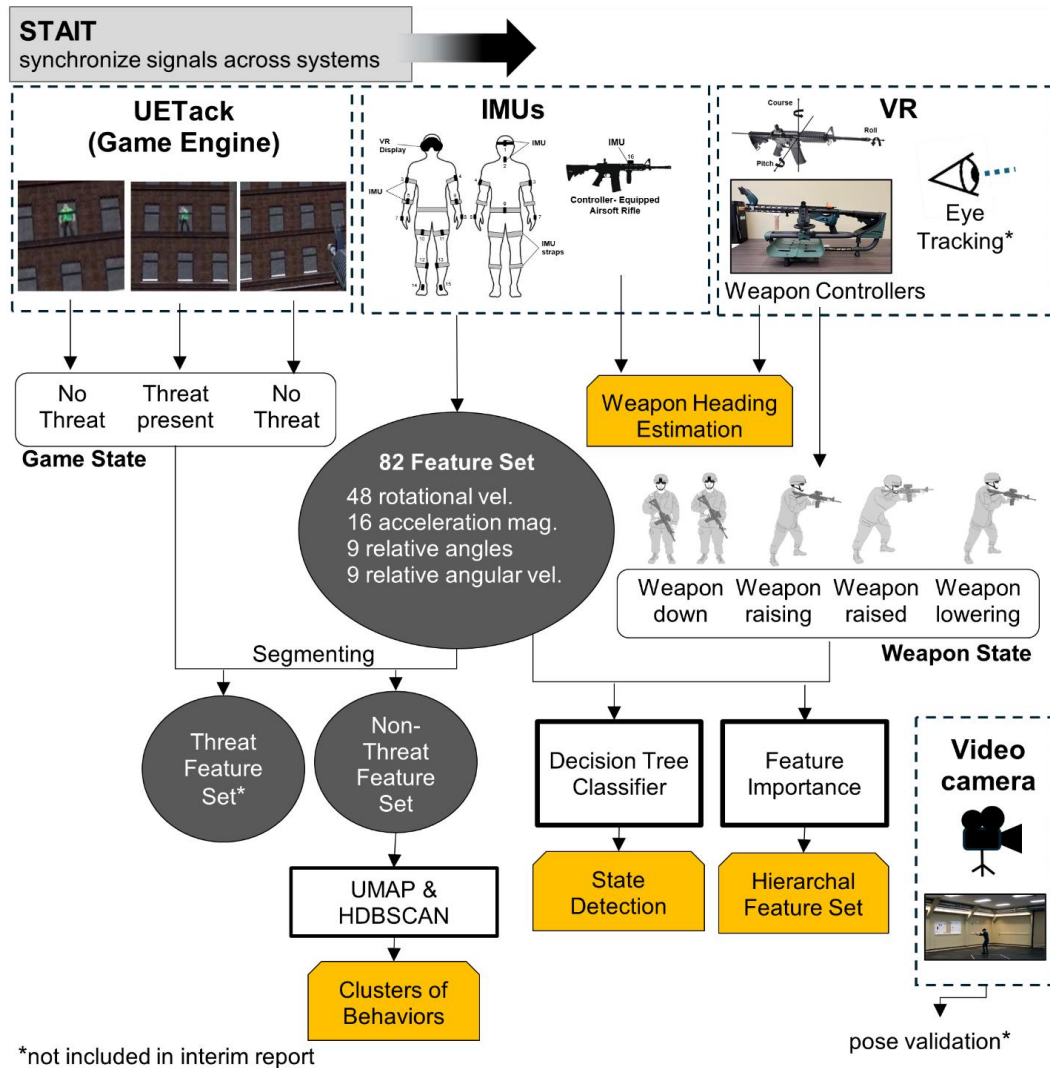


Figure 5. General overview of processing, with gold boxes showing outcomes demonstrated in this report.

2.4.2 STAIT

Raw data collected from the HMD, IMUs, virtual environment and eye tracker do not natively exist in the same spatial and temporal reference frames, as no standard coordinate system or sampling rate exists for VR devices or experimentation. To interpret these signals cohesively for STRIKE model and algorithm development, the STAIT toolbox was developed as an in-house solution. In brief, STAIT facilitates synchronization and analysis of heterogenous spatiotemporal data and is generalizable to support various algorithm development pipelines. STAIT achieves this by defining data standards and providing software tools built around the

standard to wrangle, process, and analyze data from mixed-reality environments. First, relevant data was spatially aligned to a common reference frame (in world coordinates) with position and orientation represented as 6 degrees of freedom unit vectors. Specifically, alignment was achieved by rotating and converting spatial data into a right-handed coordinate system ($x^+ = \text{forward}$, $y^+ = \text{left}$, $z^+ = \text{up}$) using centimeters as units when required. Data was then temporally aligned to a common temporal origin and converted into milliseconds (Level 0). Finally, data was interpolated to 1000 Hz regardless of original sampling rate (Level 1). This interpolation enables sample-to-sample matching across modalities for ease of analysis and visualization.

In addition, STAIT is a scalable solution to represent data from multiple individuals or agents, corresponding to their position, perspective, orientation, and environment. Importantly, because all data within STAIT conforms to the standard, the same STAIT-based algorithms for detection or inference can be applied to any dataset, thus allowing for insight into the generalizability of HMI systems and behavior. See Appendix B for a high-level overview of STAIT and the GitLab repository (Cohen Hoffing et al. 2025) for comprehensive documentation.

2.4.3 IMU Feature Extraction

Features were derived directly from the 9DoF IMU sensor data only. Although the MyoResearch MR4 recording software (Noraxon) outputs real-time joint angle estimates using state-of-the-art biomechanics models, the final SRSS suite is intended to be device agnostic. This ensures any multisensor measures could be reproduced using any IMU system and the data necessary for behavior detection is not dependent on Noraxon’s proprietary algorithms. Gyroscope data was added to the feature vector as-is with no further processing.

Three-axis accelerometer data from each sensor were combined into a single magnitude value using the l2-norm (Equation 1). This value was standardized using the decibel transform

$$M_{dB} = 20 \cdot \log_{10} \left(\sqrt{a_x^2 + a_y^2 + a_z^2} \right). \quad (1)$$

Joint angles were estimated using magnetometer data from adjacent pairs of sensors. For example, the IMU data of the forearm and upper arm were used to compute the angle of each elbow. In addition, the angles of the weapon relative to the head and the torso were computed, as these should change near the time of shot as the participant raises and aims the weapon. The magnetometer for each IMU measures the orientation of Earth’s magnetic field relative to the local coordinates of the sensor. Assuming minimal change in this field in the small distance from one

sensor to another, the differences in the angle of Earth’s magnetic field, as observed in the local coordinates of each sensor, can be used to estimate a relative angle (RA) between the two sensors.

The angle was computed as

$$\theta_{12} = \cos^{-1}(U_1 \cdot U_2) \quad (2)$$

where θ_{12} is the angle between sensors 1 and 2, and U_1, U_2 are the unit vectors for two sensors, computed as

$$U = \frac{G}{\|G\|} \quad (3)$$

where G is the magnetometer data vector $[M_x, M_y, M_z]$.

Using Equations 2 and 3, sensor pairs were used to compute RA measures listed in Table 2. In addition to these RA measures, the first-time derivative of each RA measure was computed and added to the feature vector to capture rapid changes in posture (angular velocity).

Table 2. Relative angles from nine sensor pairs (from 15 separate IMU sensors).

Measure Name	Sensor 1	Sensor 2
Head–weapon	Head	Weapon
Head–torso	Head	C7
Torso–weapon	C7	Weapon
Dominant elbow	Dominant upper arm	Dominant lower arm
Nondominant elbow	Nondominant upper arm	Nondominant lower arm
Dominant knee	Dominant thigh	Dominant lower leg
Nondominant knee	Nondominant thigh	Nondominant lower leg
Dominant ankle	Dominant lower leg	Dominant foot
Nondominant ankle	Nondominant lower leg	Nondominant foot

Overall, 82 features (referred to as the IMU feature set) were used for the hierarchal model: 48 rotational velocities (velocities of each of the 16 sensors, around each of the 3-axes), 16 acceleration magnitudes (magnitude of each sensor’s 3-axes accelerometer data), 9 RAs (RA of Table 2 sensor pairs), and 9 relative angular velocities (first time derivative of Table 2 sensor pairs). Regions of time during which the participant was not engaged in the task (time between levels, time during recalibration, etc.) were removed from all IMU sensor data and the remaining data were concatenated together. Label sets (described in Sections 2.4.4 and 2.4.5) were trimmed in this manner as well.

2.4.4 Segmentation of the IMU Feature Set and Application of Unsupervised Learning Methodology to Identify Preliminary Behaviors

To demonstrate the full IMU feature set's capability to detect distinct behaviors for general states of behaviors, the set was segmented to periods of time where no threat was active. For this, set data was aligned with recorded game state event labels from the UETack event marker (e.g., time stamp of when a threat appears in the virtual environment). Game state labels identified regions of time in which 1) no threat was present and 2) there was a threat present and attacking. These regions of time distinguished when the subject is in a more passive search phase, referred to as "surveillance" state, with no known threat versus when the subject is actively searching for a threat during a "active threat" state.

The unsupervised learning algorithm was then applied to the segmented IMU feature set via a methodology that followed work previously conducted by Hsu and Yttri (2021) who developed B-SOID, an open-source algorithm to identify and predict behaviors in 2D rat movement. This methodology employed Uniform Manifold Approximation and Projection (UMAP) for dimensionality reduction (McInnes et al. 2018) and then density-based clustering using hierarchal methodologies (Campello et al. 2013) to identify activities in discrete regions across the body. To understand which features within the IMU feature set are most relevant to any distinct clustering, the normalized value of each feature was correlated with the cluster identity and visually compared.

2.4.5 Decision Tree Classifier Model to Detect Broad Weapon States

A decision tree classifier was used to determine if it was possible to detect broader states of tactical readiness in STRIKE (e.g., the transition from states of surveillance to responding to a threat attack) with the full IMU feature set. To label these broader states of tactical readiness, weapon state labels were derived from controller data from the Meta Quest system. These weapon state labels classified regions of time during which the weapon was in one of four states: 1) weapon raising to a high ready position, 2) weapon raised in the high ready position, 3) weapon lowering to the collapsed/low ready position, and 4) weapon down in the collapsed/low ready position. Weapon positional changes were detected by calculating the angle between the weapon and the floor plane (e.g., pitch of the weapon) and the rate of angular displacement (change) of this angle. Since this sensor was not included in the IMU feature set, it provided an objective means for labeling data. Large velocity changes were identified using the course rotational rate and the direction of these movements was identified using the sign of the change. The angle of the weapon between raising and lowering events was used to

confirm the weapon was in a raised or lowered position during these times. This preliminary analysis of preshot classification accuracy provided an understanding if it were possible to detect threat engagement transition at some level of fidelity with the STRIKE IMU feature set.

2.4.6 Building a Hierarchal IMU Feature Set Using IMU Feature Importance Analysis

To understand which features were most indicative of weapon state and begin building a hierarchal IMU feature set (i.e., find the most important features), IMU feature importance was determined by using MATLAB's predictor importance. This algorithm determines how influential each feature is when determining splits in the decision tree classifier, rank-ordering the critical features and sensor locations (e.g., UMAP dimensions) in the detection of weapon state labels. For this analysis, IMU feature set data was isolated to the time around when the threat was mitigated (shot) and the top 10 largest predictors in the classifier were considered the most vital features to include in the hierarchal feature set. From this, we can begin to infer what sensors should be included in the SRSS. Values included in the predictor importance analysis were normalized within subjects.

2.4.7 Estimating Weapon Heading from the IMU Weapon Sensor

Beyond individual threat response state information, heading estimation gathered from raw IMU signals—especially from multiple sources—can yield greater information about threat numbers and positions on the battlefield. Most commercial IMU systems for pose estimation use physiological models requiring many IMUs and explicit knowledge of location on the subjects' body position then correct for error, which acts as a barrier to operational use (Kim and Lee 2022). The Attitude and Heading Reference Systems converts IMU signals to headings in the Northeast Down (NED) frame of reference to simplify modeling and reduce the number of IMUs required (Wang 2025). In particular, the Madgwick algorithm, a popular method for autonomous vehicle tracking, can generate Euclidean headings in the NED reference frame for human activity recognition (Guidolin 2021). Heading data from the weapon provides a wealth of information on threats and points of interest within the environment.

To demonstrate this usability, heading estimated from the weapon IMU was compared with the heading determined from Unreal Engine (generated from the VR controllers embedded in the instrumented rifle). For this preliminary application of STRIKE, Unreal Engine was considered the “ground truth” heading. To calculate weapon heading estimation from the weapon IMU sensor, the raw signal was shifted from its local reference frame to Euclidean headings in a NED

frame of reference. The Madgwick algorithm determines headings using magnetic north and gravity to provide the north and down headings, respectively. The algorithm uses the gyroscope, accelerometer, and magnetometer in conjunction to act as filters for error and drift on any one individual sensor (Madgwick 2010). Together, this created a heading vector oriented with its base at the operators' physical location in NED space, represented as weapon course, pitch, and roll (Figure 6). Weapon course is the angle from -180° to 180° (e.g., rotating around the upwards axis) with 0° aligned with magnetic north direction. Weapon pitch is the angle from -90° with respect to the horizontal axis (pointing the weapon straight down, toward the ground) to $+90^{\circ}$ (pointing the weapon straight up, skyward). Weapon roll is the rotation angle around the forward axis (around the barrel of the gun) from -180° to 180° with 0° marking an unrotated state.

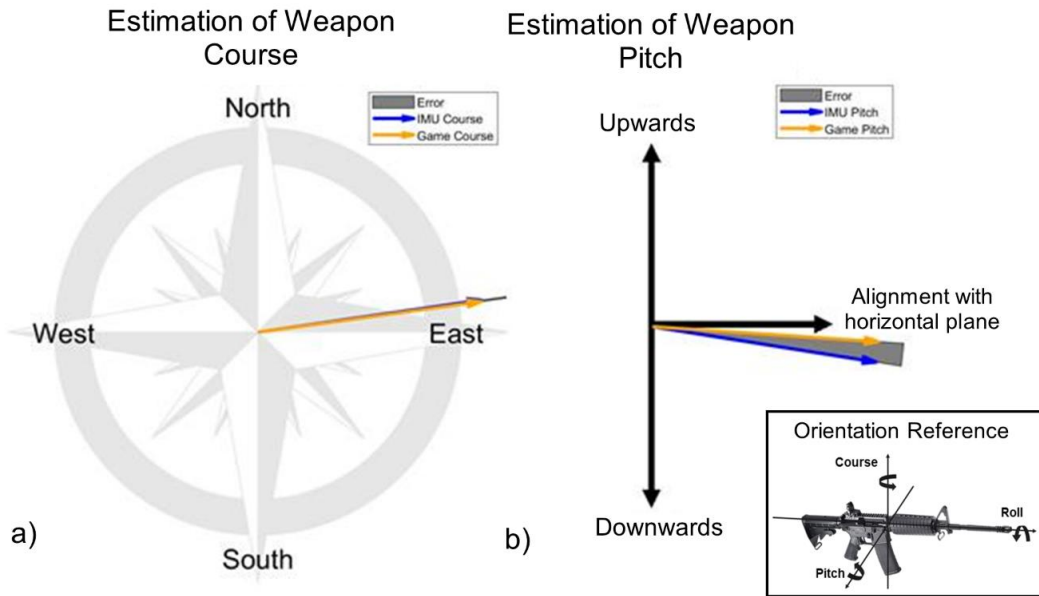


Figure 6. Data from single pilot subject shows estimated weapon heading from the IMUs compared with that from Unreal Engine controllers.

Estimated weapon heading from the weapon IMU sensor was synchronized and up-sampled to match time course of the weapon heading provided by Unreal Engine. Raw estimated weapon course, pitch, and roll were visually compared between the weapon IMU sensor and Unreal Engine and quantitatively compared by calculating the root mean square error (RMSE) between the signals. Corrections were applied to estimated IMU headings to attempt to align these with Unreal Engine headings. The correction algorithm was developed in-house and included three operations for correction. First, IMU headings were shifted by the difference between mean heading over a time segment and Unreal Engine mean. Then IMU headings were reflected across the IMU heading mean for that time sequence. Finally, IMU

headings were clipped within their accepted domains. For the pilot study, sensors were occasionally swapped out between trials. Thus, trial start and end times were used as cutoff points and operations were performed piecewise on each time segment in a loop of shifting means, reflection, and clipping outliers. Changes were kept for each operation if the RMSE error improved. Otherwise, they were rejected. For each time segment, the loop would be stopped once none of the operations had yielded an improvement in metrics. Corrected IMU headings were again compared visually with the Unreal Engine's and RMSE calculated again between the corrected weapon heading estimation and the Unreal Engine's.

3. Results

3.1 Detecting Meaningful Behaviors in Our Unsupervised Learning Methodology

Methods of unsupervised learning during the no-threat period revealed at least two distinct clusters, indicating there could be two distinct behavioral states within the no-threat condition. The 2D representation of all data points produced by UMAP is shown in Figure 7. To understand which features were relevant to the distinction of the clusters, the correlation of the normalized values of each feature with the cluster identity was calculated. The amplitude of the feature at each sample point was plotted as a color shade (red for positive, blue for negative) to better assess how feature values differ across clusters. For features that contribute more to the cluster distinction, a larger difference in feature value is expected between the two clusters. To aid in this visualization, histograms of each feature were plotted in the lower panel, showing the distribution of feature values for each cluster. Five selected features are shown as examples in Figure 7. Each feature shows distinct differences between clusters. Based on the identities of these discriminant features, it appears these two states represent stationary and moving periods, where blue color indicates a slower acceleration movement and red indicates increased acceleration in movement. Given the color distribution, the left-sided, predominantly red cluster (cluster 1) may be indicative of acceleration of movement and the right-sided, predominantly blue cluster (cluster 2) may represent a more stationary period in the no-threat, surveillance portion of the tactical readiness cycle. The way each feature contributes to each cluster is represented as “normalized feature value” (Figure 7, top row). The distribution of normalized feature values for each cluster is shown in the bottom row of Figure 7. Examination of the visualization indicates two distinct movement clusters during the nonthreat state (surveillance stage): one that represents periods of time when Soldiers were standing still or slowing down and another that may represent active movement or acceleration. Among the features

selected to be shown in Figure 7, the upper spine and dominant shank acceleration appear to be key indicators of acceleration whereas the nondominant shank, pelvis, and dominant thigh acceleration features appear to be key indicators of deceleration (among the selected variables shown here). A full examination of the full IMU feature set using this method will provide greater insight into how each feature is important to each of these clusters, providing insights into behavioral differences associated with each cluster. This analysis provides a basis to begin parsing through the features and understand their contribution to distinct clusters of movements. Focusing initially only when Soldiers were in the surveillance stage allows for a baseline comprehension of movements during tactical readiness cycle and is critical prior to adding more variable movements that occur during the threat encounter stage.

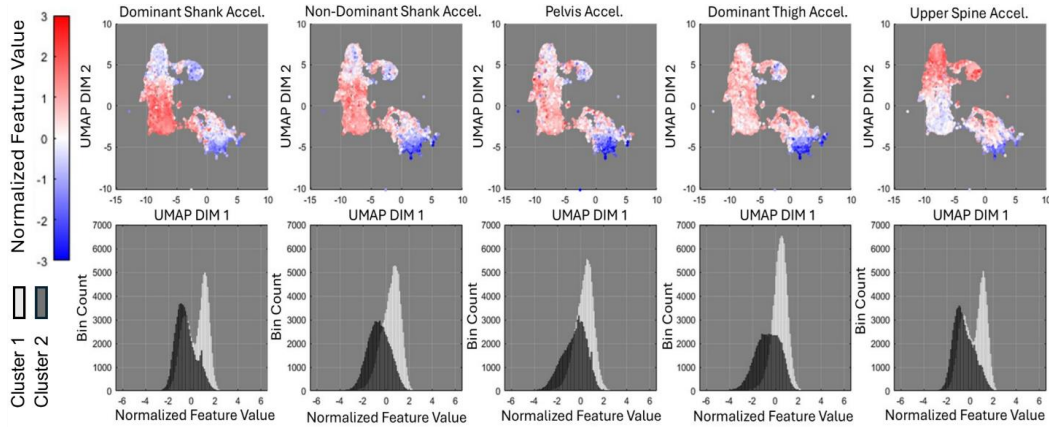


Figure 7. 2D representation of all data points produced by UMAP showing at least two distinct clusters, indicating two distinct behavioral states within the no-threat game state for a selection of five features from the IMU feature set.

3.2 Successful Detection of Broad Weapon States Using a Decision Tree Classifier Using the Full IMU Feature Set

Decision tree classification performance was assessed using a leave-one-out cross-validation scheme. On each cross-fold, the data for one subject was held out and the model was trained on data from all remaining subjects. This model was used to generate label predictions on the withheld subject's data. Performance on each cross-fold was measured using balanced accuracy (mean recall of all four classes). This measure was chosen due to its ability to provide a reliable measure of performance in cases where class sizes are greatly imbalanced. In this case, the number of samples in the weapon-down class is far greater than the other three classes. The mean balanced accuracy across all cross-folds (i.e., the four weapon states: down, rising, raised, and lowering) is $61.49\% \pm 2.69\%$ (Figure 8).

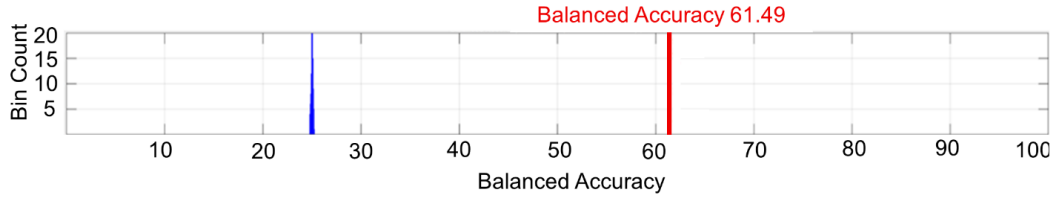


Figure 8. Decision tree classifier found above-chance accuracy when compared with accuracy of shuffled weapon state labels obtained by bootstrapping ($N = 100$).

To confirm that the model performs above chance, the null distribution of balanced accuracy was generated using a bootstrap approach. The cross-validation approach described above was repeated 100 times with labels randomly shuffled on each iteration. The distribution of mean balanced accuracy from this process is plotted shown with blue histograms in Figure 8, and the balanced accuracy obtained when using the true labels is plotted in red. The model balanced accuracy is higher than the result of every bootstrap iteration, indicating this approach identifies the four classes with above-chance accuracy.

3.3 Identification of the SRSS

Of the 82 IMU features, the predictor importance algorithm found that the 10 highest ranking features were the 9 RAs and the weapon rotational velocity about the z-axis (weapon course angle). The order of importance of these features is as follows: the nondominant elbow angle, weapon-head angle, weapon-torso angle, dominant elbow angle, nondominant knee angle, head-torso angle, dominant knee angle, nondominant ankle angle, dominant ankle angle, and then the weapon rotational velocity around the z-axis (Figure 9). Figure C-1 in Appendix C displays the predicted importance of all features.

The top 10 features were derived from sensors on the head, torso, and weapon. Given this information, the decision tree classifier was rerun with only features derived from the head, torso, and weapon sensors, resulting in an accuracy of $61.32\% \pm 2.90\%$ to detect weapon state. These results provide a strong preliminary indication that we can reasonably detect broad weapon states with a SRSS that includes only the head, torso, and weapon sensors with minimal loss in fidelity.

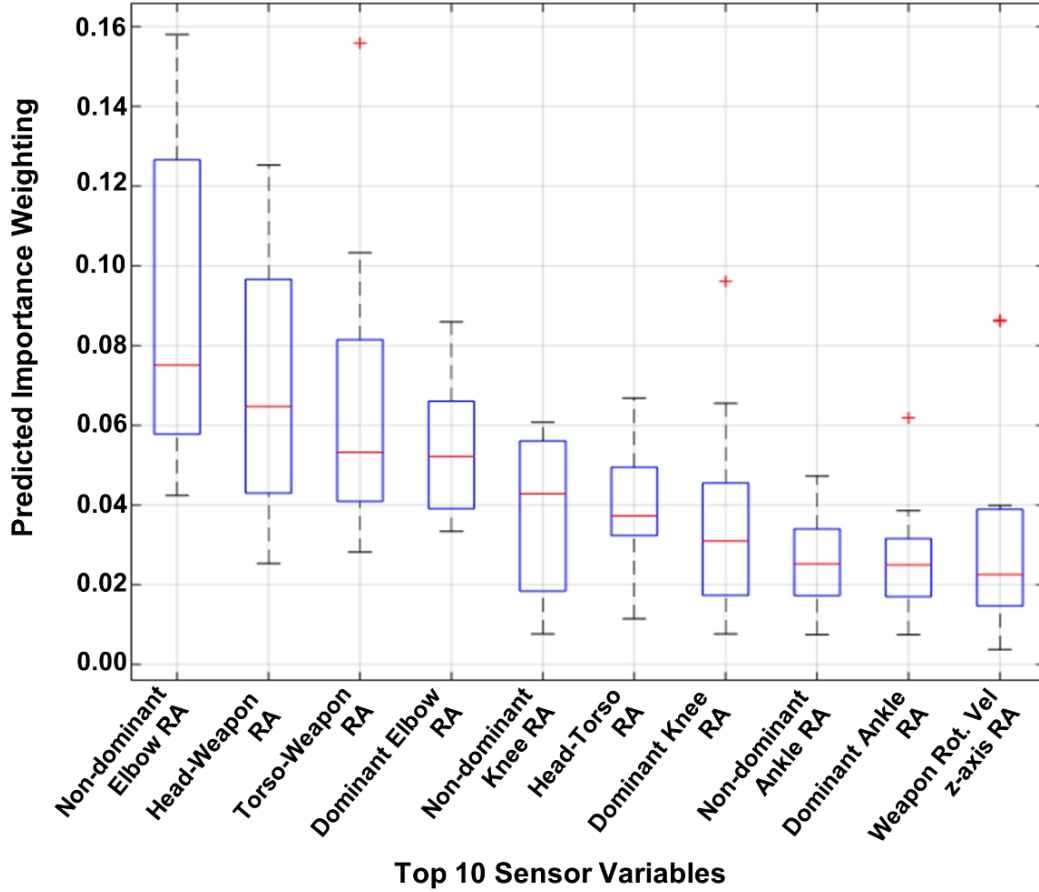


Figure 9. Ten most important features for classification of weapon state (ranked) were the RAs of nine sensor pairs (Table 2) and the weapon’s rotational velocity around the z-axis.

3.4 IMU-Estimated Weapon Heading Accuracy at Time-of-Shot and Triangulation on a Target

The capability to approximate and apply signal correction methodology for improving estimation of weapon heading directly from an IMU sensor on the weapon for one STRIKE subject is shown in Figure 10. In Figure 10, the raw, uncorrected weapon heading estimation from the IMU on the weapon is plotted against the ground truth heading provided by the VR controllers projected in the game environment. Figure 11 demonstrates how weapon heading estimation from the IMU sensor on the weapon improves with corrections applied to the signal. Signal corrections reduced the RMSE between the predictions of the IMU weapon and VR controllers by 63.8% for the weapon course, 91.7% for weapon pitch, and 67.8% for weapon roll. While roll shows significant quantitative improvement, it still appears to suffer significantly more discrepancies compared with the corrected course or pitch.

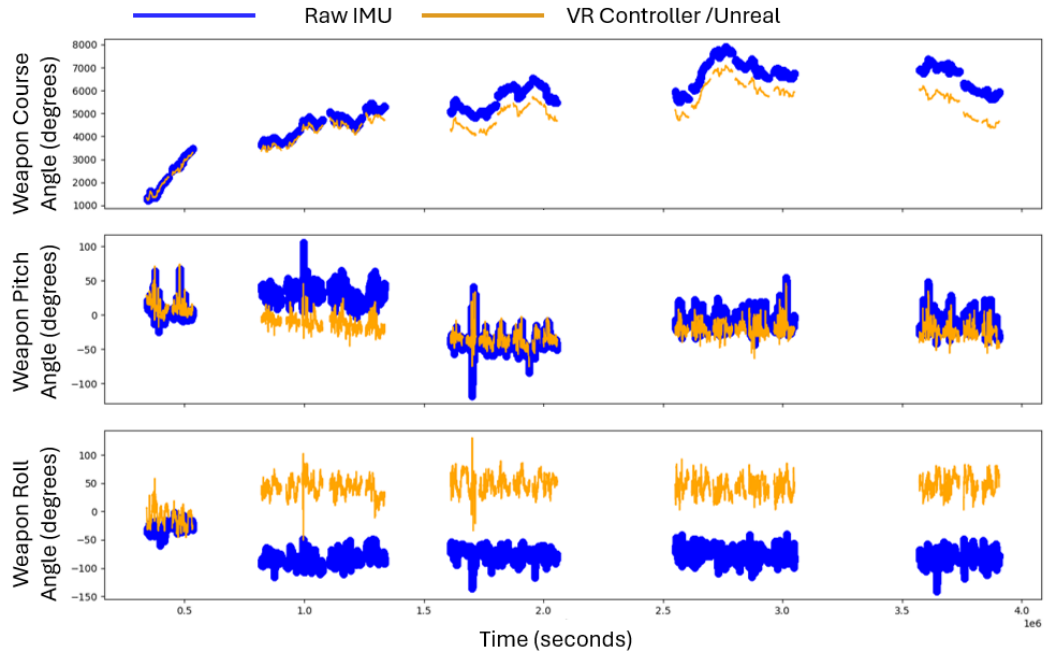


Figure 10. Raw IMU weapon heading data from a single subject, demonstrating signal discrepancies when compared with the heading estimation from the Unreal Engine and VR controller.

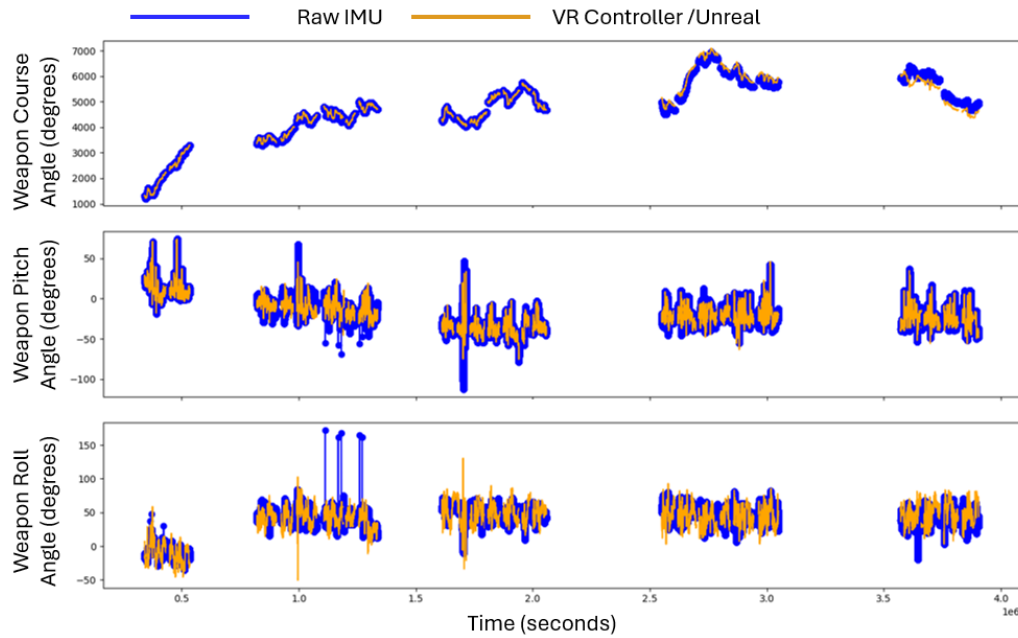


Figure 11. Correction of the IMU weapon IMU heading data improved alignment with the VR controller data from Unreal Engine for weapon course, pitch, and roll.

To demonstrate the ability to use headings to understand and pinpoint threat locations, VR controller data from two separate subjects looking at the same threat were replayed simultaneously (Figure 12). Figure 12a shows Soldier 1 (orange) and Soldier 2 (green) surveying the virtual environment and the threat is inactive. In Figures 12b through 12d, the threat becomes active and Soldiers demonstrate search with the weapon raised. In Figure 12e, both Soldier 1 and Soldier 2 have simultaneous weapon alignment with the threat, just prior (1 s) to firing upon the threat in Figure 12f. The intersection point, or vector point, of the weapon heading from multiple weapons can be used as a method of communicating points of interest in the environment via triangulation. Object depth is difficult to estimate from a single weapon heading, but when vectors from more than one weapon are used, the location of such objects can be pinpointed more closely and relayed appropriately. Figure 13 demonstrates how methods of triangulation using the vector intersection of multiple weapon headings can be used to estimate threat location. Directly prior to threat neutralization (threat shot), the error between threat location estimated from triangulation (triangulation error) drops within a few meters. Specifically, triangulation error is the difference in meters between the threat location and the estimated location by using the triangulation of two pilot subjects' weapon headings in the environment during a threat attack. Triangulation using weapon heading can be useful in indicating threat localization.

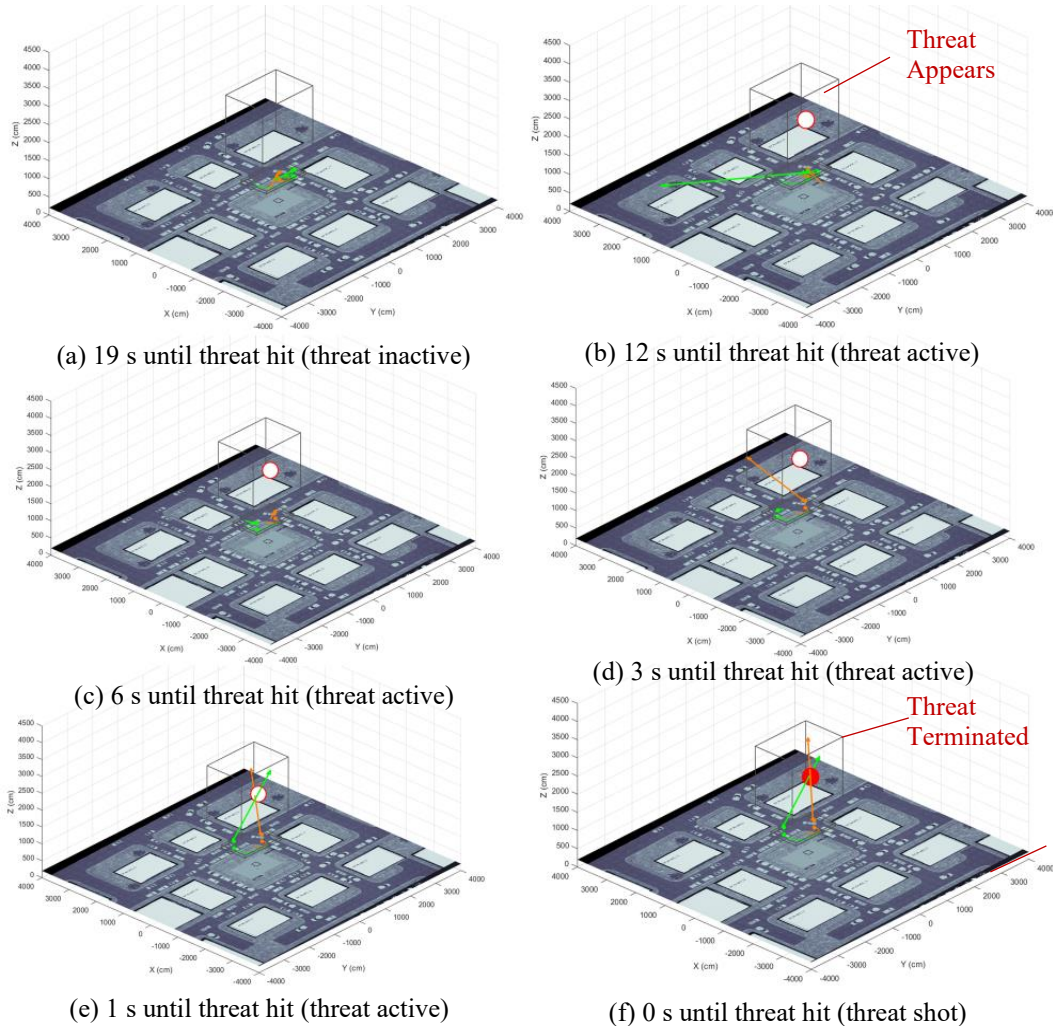


Figure 12. Estimated weapon heading from two pilot participants' Unreal Engine data of a single threat attack were simultaneously cast into the environment.

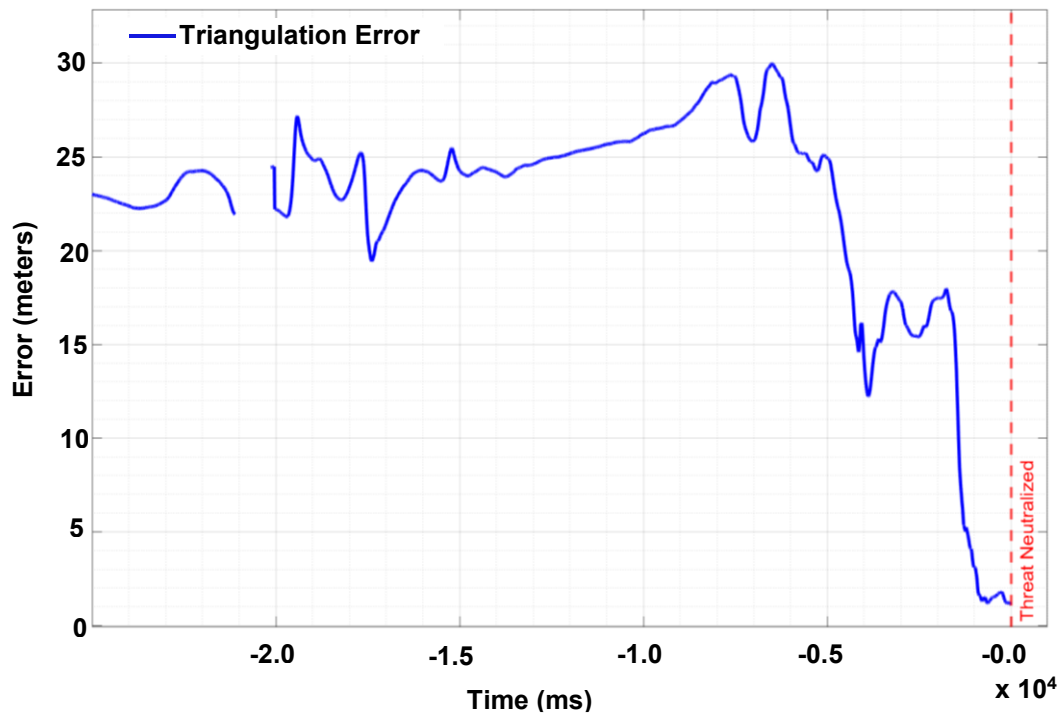


Figure 13. Triangulation error. Threat neutralized (red dash line) indicates when threat was hit (mitigated). Gaps in the data represent when one person's data was unavailable for a short amount of time.

4. Discussion

The Army's future forces must effectively integrate and employ an array of advanced weapons systems and intelligent technologies across all echelons, including at the squad and individual Soldier levels (Thompson 2022). These intelligent technologies will both require and generate an unprecedented amount of data, presenting both a challenge and an opportunity. Soldiers and commanders will find it challenging to maintain SA and make informed decisions in such a dynamic and data-rich environment, especially under disrupted, degraded, intermittent, low-bandwidth (DDIL) conditions. However, the ubiquity of these systems provides a unique opportunity to gather individual and formation data from Soldiers and robotic and autonomous systems during unit-wide tactical operations at an unprecedented level of detail.

To use this opportunity, ARL's DIESL program has focused on transforming the myriad sources of data from Soldier systems into actionable information for humans and machines without adding cognitive or physical burden through opportunistic sensing to improve small unit lethality. DIESL has developed systems and methodologies that may improve shared SA and enhance decision-making by using the Soldier and squad as a sensor to exploit natural behavior during missions

for impactful insights to improve operational outcomes. DIESL has refined a closed-loop approach of opportunistic sensing to build capabilities for curating actionable information, pushing science from controlled laboratory research to real-world validation. The STRIKE study presented within this report demonstrates this process with clear opportunities to address challenges specific to passive pointing and covert targeting via triangulation.

Preliminary analysis of the STRIKE data demonstrated reasonable capability to detect some broader tactical response states and identified some features of the model that could be most indicative of these states. The current working model with the SRSS is capable of greater than 61% accuracy in distinguishing distinct threat events in the tactical readiness cycle. Details related to the STRIKE models and algorithms will be available for others interested in detecting and interpreting meaningful behaviors with sensors during real-world maneuverers. In addition, STRIKE was the first use-case for STAIT. After further fine tuning and development, STAIT will be made open release and available for others working in this space. The following section provides some context for these results, discusses the implications, and outlines next steps in the analysis.

4.1 Building the SRSS

Our results demonstrate that data-driven exploratory analyses, such as unsupervised learning methods, can clearly identify latent behavioral states that are not directly labeled by the logged game events. This will enable further development of the label space to identify additional behaviors performed during the study. As experimentation moves to increasingly uncontrolled, self-paced environments, data-driven approaches to behavioral state identification will play an important role in interpreting data and creating behavioral identification models.

Feature importance and decision tree classifier results provide a strong preliminary indication that we can reasonably detect broad weapon states with a SRSS that includes only the head, torso, and weapon sensors (and derived features) with minimal loss in fidelity (>61% accuracy maintained). Importantly, these results support the intended application of deriving context from a combination of existing Soldier systems. Next-generation Soldier systems will likely include IMUs (or related sensors) on the weapon (smart scopes), head (heads-up display, night vision goggles), and torso (Nett Warrior system, controller).

Going forward, eye-tracking variables will be included as features in algorithm development and hierarchical models. It is expected that eye-tracking will increase accuracy of detecting weapon state, especially in discriminating stages between broader states of threat search and pinpointing the exact threat location immediately

prior to trigger pull. Additionally, we will validate the STRIKE final model using annotated video data and STRIKE models and algorithms will be optimized for real-time interpretation of behaviors during real-world field maneuvers. While the game and position data captured by the VR system provide a proxy of ground truth for model development, annotated video data remains the best way to precisely identify response states and transitions. In addition, exploratory analysis will investigate the relative timing of behavior across the body during a tactical response cycle and we will investigate how threat type (informed vs. uninformed threats) and tactical weapons expertise impacted movement behavior and response time. Such information could be used as a baseline and provide valuable information on how new training and integrated technologies impact the timing of threat detection and threat mitigation during in-field tactics.

4.2 Weapon Heading Estimates

In general, estimated weapon course and pitch angles appear to be well aligned to the VR-derived angles after adaptive processing, though some discrepancies (potentially attributed to slight differences in alignment or internal data processing between the game and IMU environments) remain. It is possible that the remaining difference could be reduced by applying conventional filtering techniques such as Fast Fourier Transform analysis and Savitsky–Golay smoothing to rectify the signal for individual error points. Weapon roll appeared to be less successfully matched to VR controller representation in the Unreal Engine, even with correction. However, use of roll will be minimal in STRIKE response state classification models, and weapon roll inclusion is currently limited in building STRIKE heading estimation models. Additionally, previous issues have been reported in the Unreal Engine-based software packages regarding the approximation of weapon roll, contributing to discrepancies. Estimation of weapon heading with weapon IMU data demonstrates the capability of these sensors to determine where the weapon is pointing in the field. Further work will include applying the filtering algorithms to improve estimations in weapon course and pitch.

Although participants were run individually within the STRIKE VR environment, post-analysis allows for the integration of data (e.g., weapon heading) from multiple people to be cast simultaneously in the VR environment. Although this heading estimation used data from the VR controller and not the IMU, it demonstrates how weapon heading from multiple weapons sensed opportunistically could be used to triangulate and identify threats or areas and points of interest. Triangulation data could also provide information on threat counts or environmental information. For instance, a large triangulation error that persists during a known threat scenario could be indicative of 1) unknown threat location

or poor line of sight for one or both Soldiers or 2) multiple threats. Such information paired with other sensor data could be useful in designing in-field support technology (e.g., unmanned aerial vehicles). This data enables the development of passive pointing and passive threat identification and localization solutions for use across the squad (e.g., by relaying threat information to other Soldiers and robots in the unit) or by higher echelons for battle tracking and fires integration. By leveraging weapon state, position, and heading estimation from multiple Soldier's weapons as they move through the VR environment and converge on a single threat, we can pinpoint visual targets and provide specific location information without electromagnetic emissions.

4.3 Army and Future Applications

The modern Army Soldier has adopted the “every Soldier is a sensor” philosophy where, in fields of operations, Soldiers and leaders are critical as information collectors (HDA 2016b), transmitting unprecedented amounts of data from Soldier-borne sensors (electro-optic, IR, acoustic, inertial, location, etc.) to the tactical network. Across the Portfolio Acquisition Executive Ground Maneuver (previously Program Executive Office Soldier), there are a broad range of programs seeking to increase Soldier lethality and survivability via these systems (and advanced gear technology), often at the cost of additional size, weight, and power that the Soldier must bear. However, data collected from sensors needs to be efficiently inferred for it to become tangible, actionable, and meaningful information, especially for seamless integration with autonomous systems without overburdening Soldiers.

Currently there are several Army initiatives working to address these challenges. An emerging priority at the Maneuver centers is developing capabilities that are passive alternatives to existing active capabilities as the threat of near-peer adversaries become likely. These passive capabilities often require multiple sensor modalities from multiple sources (i.e., Soldiers) and rely on processing and robust HMI for proper function. Current Army initiatives by the SL CFT seek to enable federated or collaborative fires across very low tactical echelons (as low as the infantry company), whereby targets sensed by any sensor in the unit can relay targeting data via tactical inter-Soldier networks to all combat-role personnel in the unit to reduce the duration of the kill chain required to service those targets with effects. Realizing this desired capability will require that tactical data be collected opportunistically from Soldiers and robots as they engage in close combat, execute reconnaissance missions, and encounter reaction forces. Related initiatives, such as Army Futures Command's HMI Formations initiative call for deploying robotic and autonomous systems within the dismounted infantry squad for a variety of tasks; these robots will benefit from human-generated data beyond what can be

sensed and perceived using on-platform compute to inform machine decision-making. Finally, efforts to integrate AI and automation into higher echelon operations (for example, the Army's Next Gen Command and Control initiative) require data perception pipelines that include edge processing and labeling to function effectively in DDIL environments. As the Army continues to invest in data-hungry technologies, it will require novel acquisition and processing approaches to turn that data into information without resorting to brute force which would be computationally intractable.

STRIKE adds to these initiatives as a foundational test-bed resource for understanding how Soldier data can be integrated with technology to improve performance in realistic field dismount maneuvers requiring integration with intelligent future systems (e.g., smart scopes; unmanned vehicles; sensors for detecting, identifying, localizing and reporting; artificial intelligence; and threat recognition technology). Additionally, STRIKE aims to identify and use the necessary amount of sensing and processing at system and subsystem levels to minimize cognitive and physical burden on the dismounted Soldier. Algorithms developed with STRIKE will improve AI integration and development, particularly in learning and planning for typical scenarios and environments. Additionally, the algorithms will provide real-time information on SA-driven movement dynamics in complex environments, especially when the SRSS is deployed on multiple teammates (e.g., team formations, number of target(s) and distractor(s), identifying areas of interest). Understanding indicators of the entire threat response cycle across the team, beyond just collecting information on weapon shots fired, provides a more holistic understanding of the threat environment such as threat number and threat locations, respective Soldier location(s), field cover, and obstacles. This holistic view of the environment could provide critical information to intelligent systems, command, and support teams for rapid and effective decision-making and allocation of assets on the battlefield.

5. Conclusion

The DIESL program is developing opportunistic sensing approaches to address future Army challenges related to the increase and rapid advancement of human-system integration into our future forces. With STRIKE, we demonstrate a successful preliminary closed-loop approach using a battery of kinematic sensors and methods of unsupervised learning to opportunistically detect behaviors associated with stages of tactical readiness. We identified an SRSS and achieved greater than 61% accuracy in detecting broad states of tactical response. We can practically implement this sensor system into existing and next-generation Soldier systems without overburdening the Soldier. In addition to disseminating critical

information across the squad in real time, this capability can provide commanders with critical threat information in dynamic battlefield environments and inform autonomous systems' development and integration. Additionally, algorithms developed from this study can inform passive pointing and covert targeting approaches, maximizing target identification and localization while minimizing detection by the enemy. Ongoing work focuses on integrating eye-tracking data to improve the model, validation with captured video data, and investigating the impact of threat type and weapon firing expertise on performance. This work may enhance shared SA and improve decision-making capabilities across the battlefield, contributing to more effective, data-driven operations.

6. References

- Ahmidi N, Ishii M, Fichtinger G, Gallia GL, Hager GD. An objective and automated method for assessing surgical skill in endoscopic sinus surgery using eye-tracking and tool-motion data. *Int Forum Allergy Rhinol.* 2012;2(6):507–515. <https://doi.org/10.1002/alr.21053>
- Aivar MP, Li C-L, Tong MH, Kit DM, Hayhoe MM. Knowing where to go: spatial memory guides eye and body movements in a naturalistic visual search task. *J Vision.* 2024;24(9):1. <https://doi.org/10.1167/jov.24.9.1>
- Campello RJ, Moulavi D, Sander J. Density-based clustering based on hierarchical density estimates. In *Pacific-Asia conference on knowledge discovery and data mining*; Gold Coast, Australia; 2013 Apr 14–17. p. 160–172. https://doi.org/10.1007/978-3-642-37456-2_14
- Cartmill EA, Beilock S, Goldin-Meadow S. A word in the hand: action, gesture and mental representation in humans and non-human primates. *Philos Trans R Soc B Biol Sci.* 2012;367(1585):129–143. <https://doi.org/10.1098/rstb.2011.0162>
- Castner N, Frankemölle J, Keutel C, Huettig F, Kasneci E. LSTMS can distinguish dental expert saccade behavior with high” plaque-urracy”. In *2022 Symposium on Eye Tracking Research and Applications*; Seattle, WA; 2022 June 8–11. p. 1–7. <https://doi.org/10.1145/3517031.3529631>
- Chen JY, Barnes MJ. Human–agent teaming for multirobot control: a review of human factors issues. *IEEE Trans Hum-Mach Syst.* 2014;44(1):13–29. <https://doi.org/10.1109/thms.2013.2293535>
- Cohen Hoffing R, Nonte M, Dang T, Touryan J. Spatiotemporal Analysis and Inference Toolkit. Army Research Laboratory (US); 2025. <https://gitlab.sitcore.net/VisualSearch/sg-stait>
- Enders LR et al. Evidence of elevated situational awareness for active duty soldiers during navigation of a virtual environment. *Plos one.* 2024;19(5):e0298867. <https://doi.org/10.1371/journal.pone.0298867>
- Enders LR, Smith R, Gordon SM, Ries AJ, Touryan J. Gaze behavior during navigation and visual search of an open-world virtual environment. *Front Psychol.* 2021;12:681042. <https://doi.org/10.3389/fpsyg.2021.681042>

- Endsley MR. Design and evaluation for situation awareness enhancement. In Proceedings of the Human Factors Society annual meeting. 1988;32(2):97–101. Sage CA: Los Angeles, CA: Sage Publications. <https://doi.org/10.1177/154193128803200221>
- Endsley MR. Toward a theory of situation awareness in dynamic systems. *Hum Factors*. 1995;37(1):85–104. <https://doi.org/10.1518/00187209577904954>
- Endsley MR, Jones DG. Situation awareness oriented design: review and future directions. *Int J Hum-Comput Int*. 2024;40(7):1487–1504. <https://doi.org/10.1080/10447318.2024.2318884>
- Fox SE, Faulkner-Jones BE. Eye-tracking in the study of visual expertise: Methodology and approaches in medicine. *Frontline Learn Res*. 2017;5(3):43–54. <https://doi.org/10.14786/flr.v5i3.258>
- Ghiani A, Van Hout LR, Driessen JG, Brenner E. Where do people look when walking up and down familiar staircases? *J Vision*. 2023;23(1):7. <https://doi.org/10.1167/jov.23.1.7>
- Giles GE et al. Characterizing relationships among the cognitive, physical, social-emotional, and health-related traits of military personnel. *Mil Med*. 2023;188(7–8):e2275–e2283. <https://doi.org/10.1093/milmed/usad002>
- Grasso R, Prévost P, Ivanenko YP, Berthoz A. Eye-head coordination for the steering of locomotion in humans: an anticipatory synergy. *Neurosci Lett*. 1998;253(2):115–118. [https://doi.org/10.1016/S0304-3940\(98\)00625-9](https://doi.org/10.1016/S0304-3940(98)00625-9)
- Guidolin M et al. On the accuracy of imus for human motion tracking: a comparative evaluation. In 2021 IEEE International Conference on Mechatronics (ICM); 2021 Mar. p. 1–6. IEEE. <https://doi.org/10.1109/ICM46511.2021.9385684>
- Halilaj E et al. Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities. *J Biomech*. 2018;81:1–11. <https://doi.org/10.1016/j.jbiomech.2018.09.009>
- [HDA] Headquarters, Department of the Army. Rifle and carbine. Headquarters, Department of the Army (US); 2016a. Training Circular No.: TC 3-22.9 https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/ARN19927_TC_3-22x9_C3_FINAL_WEB.pdf

- [HDA] Headquarters, Department of the Army. Techniques for information collection during operations among population. Headquarters, Department of the Army (US); 2016b. Army Techniques Publication No.: ATP 3-55.4. https://armypubs.army.mil/epubs/DR_pubs/DR_a/ARN43768-ATP_3-55.4-001-WEB-3.pdf
- Hessels RS, Niehorster DC, Holleman GA, Benjamins JS, Hooge ITC. Wearable technology for “real-world research”: realistic or not? *Perception*. 2020; 49(6):611–615. <https://doi.org/10.1177/0301006620928324>
- Hollands MA, Patla AE, Vickers JN. “Look where you’re going!”: gaze behaviour associated with maintaining and changing the direction of locomotion. *Exp Brain Res*. 2002;143(2):221–230. <https://doi.org/10.1007/s00221-001-0983-7>
- Hollinger D, Pollard RS, Schall MC Jr, Chen H, Zabala M. A hierarchical-based learning approach for multi-action intent recognition. *Sensors*. 2024;24(23):7857. <https://doi.org/10.3390/s24237857>
- Horst F, Lapuschkin S, Samek W, Müller KR, Schöllhorn WI. Explaining the unique nature of individual gait patterns with deep learning. *Sci Rep*. 2019;9(1):2391. <https://doi.org/10.1038/s41598-019-38748-8>
- Howe PR. Leadership and training for the fight: a few thoughts on leadership and training from a former special operations soldier. AuthorHouse; 2005.
- Hsu AI, Yttri EA. B-SOiD, an open-source unsupervised algorithm for identification and fast prediction of behaviors. *Nat Communicat*. 2021;12(1):5188. <https://doi.org/10.1038/s41467-021-25420-x>
- Hu Z, Bulling A, Li S, Wang G. EHTask: recognizing user tasks from eye and head movements in immersive virtual reality. *IEEE Trans Vis Comput Graph*. 2023;29(4):1992–2004. <https://doi.org/10.1109/TVCG.2021.3138902>
- Imai T, Moore ST, Raphan T, Cohen B. Interaction of the body, head, and eyes during walking and turning. *Exp Brain Res*. 2001;136:1–18. <https://doi.org/10.1007/s002210000533>
- Kehring KL. Immersive simulations for dismounted Soldier research. In: Savage-Knepshield P, Martin J, Lockett J, Allender L, editors. *Designing Soldier systems: current issues in human factors*. CRC Press; 2012. ISBN: 9781317152071
- Kim M, Lee S. Fusion poser: 3D human pose estimation using sparse IMUs and head trackers in real time. *Sensors*. 2022;22(13):4846. <https://doi.org/10.3390/s22134846>

- Kolaghassi R, Marcelli G, Sirlantzis K. Effect of gait speed on trajectory prediction using deep learning models for exoskeleton applications. *Sensors*. 2023;23(12);5687. <https://doi.org/10.3390/s23125687>
- König P et al. Eye movements as a window to cognitive processes. *J Eye Movement Res*. 2016;9(5):1–16. <https://doi.org/10.16910/jemr.9.5.3>
- Lance BJ et al. Minimizing data requirements for soldier-interactive AI/ML applications through opportunistic sensing. In: *Proc. SPIE 11413, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications II*; 2020 Apr 23; online only. SPIE; c2020. p. 1141306. <https://doi.org/10.1117/12.2564515>
- Lou B, Lu T, Raghu TS, Zhang Y. Unraveling human-AI teaming: a review and outlook (version 2). *arXiv*; 2025 Apr 8. arXiv: 2504.05755v2. <https://doi.org/10.48550/arXiv.2504.05755>
- Madgwick SO. An efficient orientation filter for inertial and inertial/magnetic sensor arrays. 2010:113–118. <https://forums.parallax.com/uploads/attachments/41167/106661.pdf>
- Mavor MP, Ross GB, Clouthier AL, Karakolis T, Graham RB. Validation of an IMU suit for military-based tasks. *Sensors*. 2020;20(15):4280. <https://doi.org/10.3390/s20154280>
- McGinnis RS et al. Validation of complementary filter based imu data fusion for tracking torso angle and rifle orientation. *Proceedings of the ASME 2014 International Mechanical Engineering Congress and Exposition*. 2014 Nov 14–20; Montreal, Canada. c2014. <https://doi.org/10.1115/IMECE2014-36909>
- McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction. *arXiv*; 2018 Feb 9. arXiv:1802.03426. <https://doi.org/10.48550/arXiv.1802.03426>
- National Academies of Sciences, Engineering, and Medicine. Human-AI teaming: state-of-the-art and research needs. The National Academies Press; 2022. <https://doi.org/10.17226/26355>
- O'Donovan MP, Hancock CL, Bode VG, Hasselquist L. A comparison of expert and novice marksmanship performance and postural mechanics using inertial measurement units (IMUs) during dynamic live-fire shooting. *Appl Ergon*. 2024;114:104131. <https://doi.org/10.1016/j.apergo.2023.104131>

- Pozzo T, Stapley PJ, Papaxanthis C. Coordination between equilibrium and hand trajectories during whole body pointing movements. *Exp Brain Res*. 2002;144(3):343–350. <https://doi.org/10.1007/s00221-002-1052-6>
- Ries AJ, Callahan-Flintoft C, Madison A, Dankovich L, Touryan J. Decoding target discriminability and time pressure using eye and head movement features in a foraging search task. *Cogn Res*. 2025;10(1):53. <https://doi.org/10.1186/s41235-025-00657-y>
- Salehi M, Javadpour N, Beisner B, Sanaei M, Gilbert SB. Cybersickness detection through head movement patterns: a promising approach. *Proceedings of the 5th International Conference on Artificial Intelligence in HCI*. 2024 Jun 29–July 4; Washington, DC. c2024. p 239–254. https://doi.org/10.1007/978-3-031-60611-3_18
- Siposova B, Tomasello M, Carpenter M. Communicative eye contact signals a commitment to cooperate for young children. *Cognition*. 2018;179:192–201. <https://doi.org/10.1016/j.cognition.2018.06.010>
- Talarico M, Neugebauer J. Spatiotemporal and lower body kinematic differences between normal walking and walking while holding weapon aim. *Journal of Science and Medicine in Sport*. 2017;20:S174. <https://doi.org/10.1016/j.jsams.2017.09.585>
- Thompson M. Soldier Lethality team reimagines movement, vision and combat capabilities. U.S. Army. 2022 Mar 15. https://www.army.mil/article/254732/Soldier_lethality_team_reimagines_movement_vision_and_combat_capabilities
- Touryan J et al. UETack: Software Tools for Multiplayer Experimentation Using the Unreal Engine. DEVCOM Army Research Laboratory (US); 2023. Report No.: ARL-TR-9646. [Report is CUI]
- Wang H. Human joint motion data capture and fusion based on wearable sensors. *Autonom Intell Syst*. 2025;5(1):12. <https://doi.org/10.1007/s43684-025-00098-w>
- Zago M, Kleiner AFR, Federolf PA. Machine learning approaches to human movement analysis. *Front Bioeng Biotechnol*. 2021;8:638793. <https://doi.org/10.3389/fbioe.2020.638793>
- Zaroug A, Lai DT, Mudie K, Begg R. Lower limb kinematics trajectory prediction using long short-term memory neural networks. *Front Bioeng Biotechnol*. 2020;8:362. <https://doi.org/10.3389/fbioe.2020.00362>

Zhang T et al. Physiological measurements of situation awareness: a systematic review. *Hum Factors*. 2023;65(5):737–758. <https://doi.org/10.1177/00187208209690>

Appendix A. The Six Stages of Tactical Movement

The following outlines the stages of tactical movement in greater detail, including postural differences to potentially distinguish between stages.

- 1) Stage 1 (low threat) is when Soldiers keep the weapon at a low hanging position or collapsed low ready position with the firing hand on the weapon's grip, with weapon slung across the chest towards the nonfiring hand, and nonfiring hand on the weapon's hand guards or vertical grip. They maintain control of the weapon orientation during movement, a stage or position used when there is likely no immediate threat in the environment. It indicates a low engagement level of tactical response. The Soldier will walk with normal gait and full upper- and lower-body rotation and movement. Head and gaze movements will include lower angular velocities and large angular displacement, demonstrating a free search pattern of behavior.
- 2) Stage 2 (threat alert) is when contact with a threat is likely and the weapon is positioned at the low ready (or high ready where appropriate). Low ready upper-body position has the rifle with butt stock against the shoulder, firing hand on the grip with finger extended out of trigger guard, and nonfiring hand on weapon's hand guards or vertical grip. The weapon is held at the low ready position with muzzle elevation angle at 30° – 45° below horizon. For a situation where the alternative high ready position is more appropriate, the weapon hold is similar to low ready except the butt stock is in the armpit and the muzzle is 45° above horizon. The threat area is $\pm 45^{\circ}$ in the azimuth for both low and high ready stances. For either low or high ready, the Soldier can rapidly engage with a threat with minimal upper-body movement. Soldier focus would be forward with reduced SA to the posterior position. Movement of the trunk or lower body is such that firing stability is maximized by rolling of the feet while walking, maintaining feet shoulder width apart, and employing an aggressive lean where center of mass (trunk) shifts forward. Head and gaze movements will include faster angular velocities and angular displacement, demonstrating a directed search pattern of behavior.
- 3) Stage 3 (threat detected) is when a threat is detected, confirmed to be in the general area, but exact location is unknown. The weapon would be in the ready (or ready up) position with the butt of the rifle stock tucked into the shoulder and muzzle aligned in the most likely direction of the threat location. Placement of the hands on the weapon and lower body movements are similar to Stage 2 (threat alert). The stock weld is stabilized against the Soldier's head, with gaze through the weapon sights or optics, tracking the threat. The trunk and rifle are most likely maneuvering in a turret-like motion, with rifle and torso moving in sync. Head and gaze movements will

include faster angular velocities and angular displacement, demonstrating a directed search pattern of behavior to find the exact threat location.

- 4) Stage 4 (threat identified) is when threat is detected, exact location is fixed upon, threat is identified, and threat engagement is imminent. Similar to stage 3 (threat detection), the weapon is in the ready position, with stock weld stabilized, gaze through weapon sights, and weapon fixed on the threat. Hand and trigger finger placement on the weapon and lower body (movement) is similar to stages 2 and 3. Forward progression towards the threat may continue but head, eye, trunk, and weapon rotational velocities and transverse movement will be minimal. Gaze fixation will be locked onto the threat.
- 5) Stage 5 (threat engaged) is when lethal force has been decided, the weapon is aimed at the threat, finger is placed on the weapon trigger, final weapon adjustments are made, and the threat is engaged by pulling the trigger. This stage is the same or similar to stage 4 (threat identified) in terms of head, gaze, and weapon movement but the finger moves to the trigger and trigger pull action occurs once threat confirmation occurs. Forward movement toward the threat may occur or feet may be planted to account for stabilizing the body for weapon recoil. Small adjustments in weapon position may occur if more than one shot occurs for a specific threat.
- 6) Stage 6 (threat confirmed) is in the immediate aftermath of a threat when it is confirmed to be mitigated. The finger is removed from the trigger well and the Soldier returns to the appropriate stage from 1 through 4 depending on threat levels and intel. Head, gaze, and weapon movement will increase in angular and forward movement, and forward progression movement may begin again if the Soldier stopped movement during stage 5 (threat engaged).

Appendix B. Spatiotemporal Analysis and Inference Toolkit (STAIT) Overview

B.1 STAIT Toolbox Description and Usability

Immersive technologies like virtual, augmented, and mixed realities (MRs) are revolutionizing industries, enabling new ways to interact with digital environments. However, the data generated by these technologies—spanning diverse devices, sensors, and formats—poses significant challenges for researchers seeking to analyze human and agent behavior. Existing tools in the MR community provide a foundation, but none fully address the complexities of multiagent data, MR head-mounted displays, and eye trackers, which often involve 3D and 2D coordinates, images, and inertial measurement units.

The STAIT is a code repository available on GitLab that is designed to overcome these challenges by standardizing spatiotemporal data and providing a user-friendly platform for analysis. It empowers researchers to study human and agent behavior, interactions, and movements in 3D spaces over time, unlocking new possibilities for understanding behavior in MR environments and developing novel capabilities.

At its core, the STAIT toolbox helps researchers understand three fundamental aspects of behavior within 3D spaces:

- 1) Where someone or something is located in a 3D environment
- 2) What they are looking at or interacting with
- 3) When did these actions or interactions occur

B.2 STAIT Toolbox Key Features

- Data standardization: STAIT aligns data from diverse devices and sensors to a common format, facilitating comparison and analysis.
- Generalized algorithm development: researchers can apply and test algorithms across different datasets, regardless of the source device or sensor.
- Data processing: built-in functions simplify tasks like filtering, resampling, and event detection for spatiotemporal data.
- Object-oriented code: modular and intuitive code design makes the toolkit accessible to users with varying levels of expertise in 3D data analysis.
- Flexible integration: users can incorporate custom algorithms or tools to suit their specific research needs.
- Behavior analysis: tools for detecting eye movements (e.g., fixations, saccades) and analyzing interactions with objects in the environment are included.
- Comprehensive documentation and tutorials: designed for the extended reality community, STAIT offers extensive resources to help users get

started and maximize its capabilities (<https://gitlab.sitcore.net/VisualSearch/sg-stait/stait-core>).

B.3 STAIT Toolbox Workflow

The STAIT pipeline guides users through the process of preparing, processing, and analyzing spatiotemporal data (Figure B-1). Comprehensive documentation ensures compatibility with the STAIT standard, while tutorials help researchers wrangle collected data into the required format. Once standardized, users can apply STAIT’s tools to process data and apply algorithms for behavioral analysis.

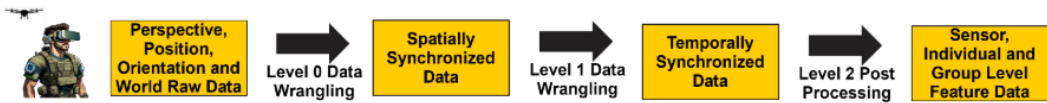


Figure B-1. STAIT pipeline.

Appendix C. Full Inertial Measurement Unit Feature Set

Importance of all features in the full Stages of Tactical Response Inferred from Kinematics and Eye-Tracking (STRIKE) feature set in classifying the broad four weapon states (weapon down, raising, up, and lowering) (Figure C-1). The top 10 most important features are shown with asterisk symbols.

Predicted Importance Weighting

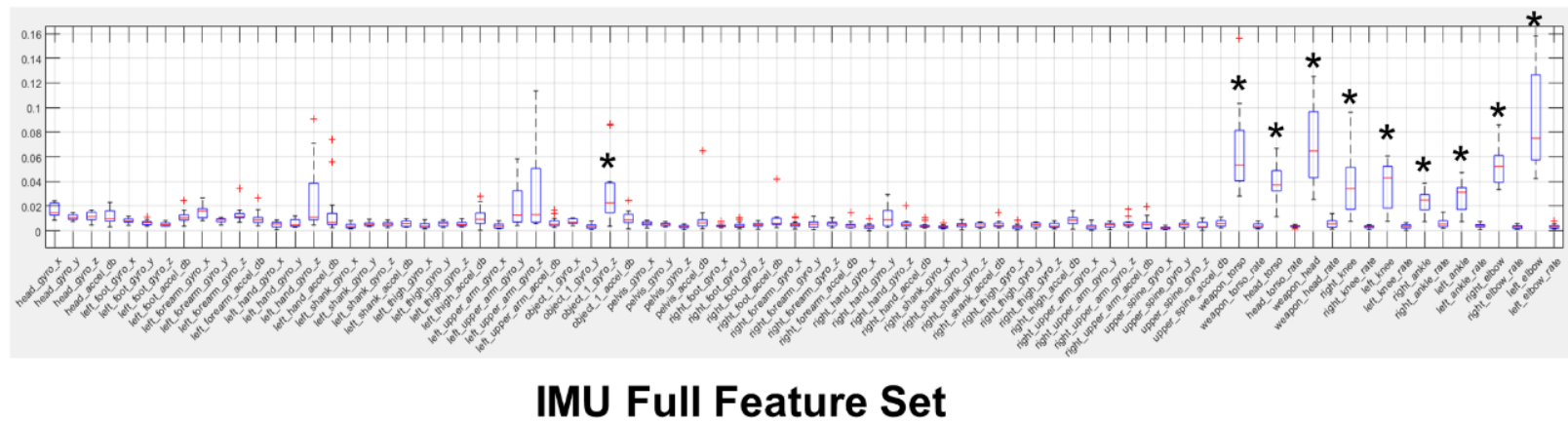


Figure C-1. Predicted importance weightings for the full set of features used in the STRIKE model. IMU = inertial measurement unit.

List of Symbols, Abbreviations, and Acronyms

2\3D	2\3-Dimensional
9DoF	9 Degrees of Freedom
AHR	Attitude and Heading Reference Systems
AI	Artificial Intelligence
ARL	Army Research Laboratory
C5ISR	Command, Control, Communications, Computers, Cyber, Intelligence, Surveillance, and Reconnaissance
DDIL	Disrupted, Degraded, Intermittent, Low-Bandwidth
DEVCOM	U.S Army Combat Capabilities Development Command
DIESL	Distributed Information for Enhanced Squad Lethality
HMD	Head-Mounted Display
HMI	Human–Machine Integration
IMU	Inertial Measurement Unit
IR	Infrared
ML	Machine Learning
MR	Mixed Reality
NED	Northeast Down
OODA	Observe, Orient, Decide, Act
PC	Personal Computer
RA	Relative Angle
RMSE	Root Mean Square Error
SA	Situational Awareness
SL CFT	Soldier Lethality Cross-Functional Team
SRSS	STRIKE Reduced Sensor Set
STAIT	Spatiotemporal Analysis and Inference Toolkit

STRIKE	Stages of Tactical Response Inferred from Kinematics and Eye-Tracking
UMAP	Uniform Manifold Approximation and Projection
VR	Virtual Reality

1 DEFENSE TECHNICAL
(PDF) INFORMATION CTR
DTIC OCA
1 DEVCOM ARL
(PDF) FCDD RLB CB
TECH LIB