NASA/TM–20250009313

Aerial AId Closeout Report

J. Tanner Slagel, Sarah M. Lehman, David Wagner, Josh Fody, Kyle Smalling, John V. Siratt, Aaron Dutle
NASA Langley Research Center, Hampton, Virginia

Nelson Brown, Ricardo Arteaga
NASA Armstrong Flight Research Center, Edwards, California

Massimo Vespignani
KBR Inc at NASA Ames Research Center Moffett Field, California

Sungshik Yim
Analytical Mechanics Associates, Hampton, Virginia

Jack Fortner-Monegan, Brandon Ruffridge
NASA Glenn Research Center, Cleveland, Ohio

Adharsh Kandula
Northeastern University, Boston, Massachusetts

Kyler Shu
Stanford University, Stanford, California

Garrett Wright
Virginia Polytechnic and State University, Blacksburg, Virginia

Owenn Herrmann
Iowa State University, Ames, Iowa

November 2025

NASA STI Program Report Series

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI Program provides access to the NTRS Registered and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI Program, see the following:

- Access the NASA STI program home page at `http://www.sti.nasa.gov`

- Help desk contact information: `https://www.sti.nasa.gov/sti-contact-form/` and select the "General" help request type.

NASA/TM–20250009313

Aerial AId Closeout Report

J. Tanner Slagel, Sarah M. Lehman, David Wagner, Josh Fody, Kyle Smalling, John V. Siratt, Aaron Dutle
NASA Langley Research Center, Hampton, Virginia

Nelson Brown, Ricardo Arteaga
NASA Armstrong Flight Research Center, Edwards, California

Massimo Vespignani
KBR Inc at NASA Ames Research Center Moffett Field, California

Sungshik Yim
Analytical Mechanics Associates, Hampton, Virginia

Jack Fortner-Monegan, Brandon Ruffridge
NASA Glenn Research Center, Cleveland, Ohio

Adharsh Kandula
Northeastern University, Boston, Massachusetts

Kyler Shu
Stanford University, Stanford, California

Garrett Wright
Virginia Polytechnic and State University, Blacksburg, Virginia

Owenn Herrmann
Iowa State University, Ames, Iowa

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23681-2199

November 2025

Acknowledgments

Aerial AId was the result of the Convergent Aeronautic Solutions discovery process that engaged hundreds of NASA civil servants, scientists, and citizens of the United States; each of them are to thank for the resulting activity and progress made. Specifically, the Aerial AId team would like to thank the dispatchers, emergency medical technicians, fire chiefs, and UAS pilots for providing their experiences that shaped this effort.
Additionally, there are several former and current NASA employees and contractors who were instrumental in the inception of Aerial AId. This includes Rachel Moore Best, Kathleen Bond, Eric Brubaker, and Elizabeth Reiken. Thank you all for your contributions.
This work was supported by the NASA Aeronautics Research Mission Directorate, through the Convergent Aeronautics Solutions (CAS) Project.

Generative AI was employed in the compilation and organization of this report and in draft writing; however, all research is augmented and fact-checked by a human being. The final draft of the report is written, edited, and proof-read solely by human subject matter experts.

Available from:

NASA STI Program / Mail Stop 150
NASA Langley Research Center
Hampton, VA 23681-2199

Abstract

This report provides an overview of the Aerial AId activity, which focused on enabling assured computer vision onboard Uncrewed Aerial Systems (UAS) to deliver rapid situational awareness for emergency response operations. Conducted under the Convergent Aeronautics Solutions (CAS) project within NASA's Aeronautics Research Mission Directorate, Aerial AId emerged from the CAS Discovery process in fiscal year 2024 and was executed through fiscal year 2025. The project aimed to address capability gaps in Drone-as-a-First-Responder (DFR) operations by developing and assuring trustworthy AI/ML perception engines tailored for aerial medical emergency scene assessment. Key accomplishments included the development of a runtime assurance framework for computer vision, the creation of a machine learning dataset focused on human stance detection for medical UAS applications, and the training of prototype object detection algorithms using aerial imagery. Aerial AId advanced the state of trustworthy computer vision technologies for autonomous DFR operations, contributing capabilities that support NASA's mission to develop aeronautics technologies for societal benefit. This report documents the project's formation, planning and execution, technical approach and outcomes, stakeholder engagement, and lessons learned—offering recommendations for future development of emergency response UAS technologies.

Contents

List of Figures

Part I
Introduction & Background

1 Introduction

The Aerial AId activity represented an effort within the National Aeronautics and Space Administration's (NASA) Convergent Aeronautics Solutions (CAS) project, [1], to enable and assure computer vision from an aerial perspective in support of a fully autonomous drone-as-a-first-responder (DFR) concept. Emergency medical responders currently operate with limited situational awareness of incident scenes before arrival, forcing them to make decisions about time, personnel, and resource allocation with incomplete information about the location, severity, and hazards present on the scene. This fundamental challenge impacts response effectiveness, including the safety of both first responders and patients. The Aerial AId project emerged from extensive stakeholder engagement through the CAS Discovery Process, which identified assured machine-enhanced situational awareness as a key technology that could dramatically improve DFR operations across most emergency medical use cases.

Executed through Fiscal Year 2025 as a collaborative effort that included NASA's Langley Research Center, Glenn Research Center, Armstrong Flight Research Center, and Ames Research Center, Aerial AId was designed to answer the fundamental research question: "Can a small uncrewed aerial system with machine vision provide useful, assured situational awareness for emergency medical response?" The project's mission was to enable a fully autonomous DFR capability by developing trustworthy computer vision technology specifically tailored for emergency medical scene assessment. This ambitious goal required addressing multiple technical and operational barriers, including the lack of relevant training datasets for aerial medical emergency scenarios, the need for assurance of safety-critical AI/ML computer vision components in emergency environments, and the requirement for onboard perception and decision-making capabilities in autonomous systems.

The project was structured around four core technical pillars that would collectively advance the state of the art in DFR systems. First, the development of application-specific datasets through flight testing, synthetic data creation, and data collection from stakeholders to support computer vision system training and validation. Second, the creation of nominal computer vision systems optimized for rapid emergency scene assessment, focusing on patient identification and stance classification. Third, the implementation of a runtime assurance (RTA) framework to monitor and ensure the reliability of AI/ML components operating in safety-critical scenarios. Fourth, the integration of the developed AI/ML component with the RTA framework was to be implemented on an edge hardware device on an Uncrewed Aircraft System (UAS), culminating in a flight demonstration of the assured computer vision system for the autonomous DFR concept. This approach was designed to provide the foundational data and technologies necessary for the commercial emergency medical UAS sector to develop safe and effective computer vision for emergency medical response applications onboard their aircraft.

2 Emergency Medical Response Challenges

The United States emergency medical services (EMS) system operates under significant strain, handling approximately 240 million emergency response calls annually through roughly 9,000 largely independent dispatch centers [2]. A critical challenge facing this system is that emergency medical responders have limited situational awareness of incident scenes before arrival, forcing them to make critical decisions and resource allocations based on incomplete or inaccurate information. Such challenges range from assessing the exact nature and severity of the call to determining the location of the emergency scene. This information gap affects every aspect of emergency response, from initial dispatch decisions to on-scene resource deployment and patient care strategies. The problem is compounded by the fact that a large percentage of all 911 calls are classified as non-emergencies, yet dispatchers lack reliable methods to accurately assess the true severity of incidents, leading to systematic over-deployment of expensive emergency resources [3]. This can result in wasted resources and a reduced level of care for individuals, which becomes exacerbated in rural areas where the response time increases due to distance traveled [4].

When inadequate situational awareness extends throughout the emergency response chain, EMS dispatchers routinely default to sending more care to a scene than may be necessary as a safety precaution. This conservative approach strains system resources and reduces availability for genuine emergencies. Conversely, there are frequent instances where scenes are worse than initially reported, requiring additional providers, medical equipment, and vehicles that were not prepositioned. As one EMS provider explained:

> *"We generally preplan based on what we're being told is occurring through the dispatch system. You may be going somewhere and are told it's a car accident, but [you arrive and] the car is upside down. You start considering, where am I gonna take this patient? What's most appropriate for them? What additional resources am I going to need? Am I gonna need extra fire department? Am I gonna need [the fire department's] ambulance?"*

This reactive approach to resource management increases response times, compromises patient outcomes, and elevates risks for both patients and first responders.

The operational challenges are further exacerbated by equipment and supply prediction difficulties. EMS providers frequently encounter situations where they lack critical medical supplies such as gauze, medications, or specialized equipment because they could not accurately predict scene requirements in advance. This forces responders to either provide suboptimal care or delay treatment while rerouting to acquire necessary supplies, both of which can jeopardize patient health. The problem is particularly acute in rural areas where resource margins are thinner and travel distances to resupply points are greater. Additionally, first responders often spend valuable time searching for patients due to imprecise location information, sometimes knocking on wrong doors and potentially compromising their own safety in the process.

Beyond immediate operational impacts, these systemic challenges create broader healthcare and economic consequences. Air ambulance pilots report that landings can be especially dangerous and that advanced understanding of emergency scenes and potential landing zones would significantly improve flight safety. Additionally, the cumulative effect of inefficient resource deployment, extended response times, and suboptimal patient outcomes contributes to healthcare cost inflation while simultaneously reducing the overall effectiveness of emergency medical services.

These interconnected challenges highlight the critical need for technological solutions that can provide reliable, real-time situational awareness to emergency medical responders before they arrive on scene. However, any such solution must be carefully designed to enhance rather than burden the cognitive load of dispatchers and EMS providers who are already operating under high-stress conditions with demanding time constraints. The complexity of this operational environment, combined with the life-or-death nature of emergency medical response, creates a unique technological challenge that requires both technical innovation and deep understanding of emergency response workflows and human factors considerations.

3 Drone as a First Responder

DFR technology represents a paradigm in emergency response that has the potential to address some of the fundamental challenges facing current emergency medical services. DFR involves dispatching drones ahead of human responders to provide real-time scene assessment before personnel arrive on location. This approach takes advantage of a UAS's ability to deploy rapidly, operate safely in hazardous environments, and provide useful aerial perspectives that are impossible to achieve from ground level. Research conducted by organizations like the MITRE Corporation has documented significant benefits of DFR programs, including faster response times, improved safety for first responders, and enhanced cost-effectiveness compared to traditional emergency response methods [5].

3.1 Current State of DFR

Currently, DFR operations in the US are primarily conducted through police municipalities and are far from autonomous operations. In the current state of DFR, a pilot flies a UAS to a scene and begins to manually survey the area. Often, a stream of the video is sent back to a station where someone visually observes the camera data coming in. A line of communication is open between the dispatcher, data stream, observer, pilot, and other relevant parties, and scene information is visually gained and verbally communicated. Beyond-Visual-Line-of-Sight (BVLOS) operations, where the UAS travels further than a visual observer can watch currently require an FAA exemption. If such an exemption is not given, then a visual observer must be assigned to watch the UAS as it is flown to the scene.

The current state of DFR operations saves lives [6], but the scalability of operations is hindered by the number of humans required per UAS being deployed. Increasing automation in the mapping and perception of the scene can ease this burden. Additionally, information gathering is currently dependent on the human eye, and information to EMS is passed primarily through audio channels, giving an opportunity for a streamlined improvement.

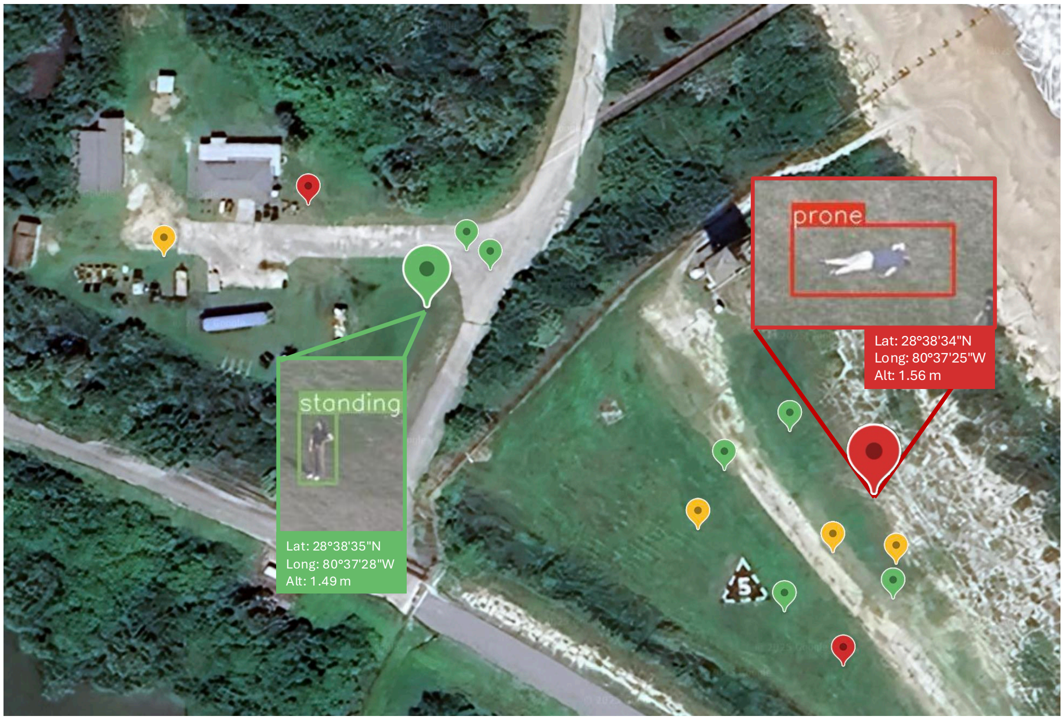

Figure 1: Overhead map visual integrating scene perception data into one interface

3.2 Autonomous DFR for Emergency Response

The vision for autonomous DFR starts with a call to dispatch. When a call comes in with an approximate location, a UAS is deployed and automatically flies to the scene. After arriving at the scene, the UAS begins to not only map the scene, but also perceive it, detecting humans, hazards, and other necessary information for safe and effective response. As it gathers scene information, a map view of the scene with relevant information begins to populate for all relevant personnel including the dispatcher, first responders on the scene, and the incident commander (in the case of large-scale disasters). Based on the information collected, the UAS might change its mapping flight path to focus on particular areas of interest. With this information, the necessary resources for the scene are determined and care is provided.

The technological foundation for autonomous DFR operations has advanced dramatically in recent years, driven by convergent improvements across multiple domains. Autonomous flight and control systems now incorporate sophisticated detect-and-avoid technologies, versatile location systems, integrated air traffic management capabilities, system failure responses, and dynamic routing algorithms that enable drones to operate with minimal human intervention. Sensor integration has reached new levels of sophistication, with modern DFR platforms capable of carrying high-definition cameras, thermal imaging systems, LiDAR sensors, and specialized detectors for gases or other hazards. Additionally, improvement of battery technology extends operational duration and coverage area. These technological advancements enable drones to capture data from emergency scenes while operating safely in dynamic and potentially dangerous environments.

From an operational perspective, DFR systems offer unprecedented capabilities for emergency scene assessment that directly address the situational awareness gaps plaguing current EMS operations. Advanced DFR platforms can provide automatic perception of emergency medical scenes, including precise location, patient identification and triage assessment, hazard detection, and environmental monitoring. The aerial perspective allows for rapid assessment of scene safety, identification of optimal ingress and egress routes, patient count and positioning, bystander locations, and environmental hazards such as fire, structural damage, or chemical spills. This intelligence can be transmitted to dispatchers and responding units in real-time, enabling informed decision-making about resource allocation, response strategies, and safety protocols before human responders enter potentially dangerous situations.

Despite the significant promise of DFR technology, substantial gaps exist between research achievements and commercial implementation, particularly in medical emergency response applications. While DFR programs are operational in law enforcement contexts, primarily for surveillance purposes, there is currently no evidence of UAS being utilized by medical first responders for scene assessment. The technology for full autonomy, especially in guidance and control systems, has seen remarkable advancements, and recent Federal Aviation Administration (FAA) regulatory updates are likely to accelerate adoption. Computer vision algorithms for fully autonomous medical emergency response scene surveillance are primed for advancement and deployment, yet their integration into commercial DFR platforms remains limited.

The barriers preventing widespread adoption of medical DFR systems include challenges related to quality assurance for autonomous systems, the lack of application-specific training datasets for computer vision algorithms, regulatory constraints, and public perception and privacy concerns. Additionally, the transition from current manned remote video surveillance approaches to fully autonomous scene assessment requires demonstrated reliability and accuracy that meets the safety-critical requirements of emergency medical response. Addressing these challenges through targeted research, improved regulatory frameworks, integration of quality assurance measures, and strategic public engagement represents the pathway forward for realizing the full potential of DFR technology in emergency medical response, ultimately enabling the vision of truly autonomous systems that can improve patient outcomes while enhancing the safety and effectiveness of first responders.

4 From Zero Impact Aviation to Aerial AId

4.1 Zero Impact Aviation Origins

Near the end of 2021, NASA's Aeronautics Research Mission Directorate ARMD began developing a white paper plan to address "Zero Emissions Aviation" - focusing on traditional aviation emissions research. The CAS Project was initially asked to assist with this effort. As the CAS team began contributing content that emphasized system-level interactions, a "discovery sprint" was launched, focusing on Zero Impact Aviation (ZIA). ZIA sought to "flip the script" on aviation sustainability research by moving beyond simply reducing negative impacts to actively identifying how aviation could be leveraged for positive societal gain[1]. The scope of ZIA encompassed end-to-end life cycle impacts of aviation across two major domains: the natural world and human society.

The ZIA sprint focused on mapping these topic areas with five major process tasks: scoping Zero Impact Aviation, developing understanding of the topic area landscape, building vignettes to envision future scenarios, identifying and evaluating areas of focus, and formulating problem spaces while gathering information to support follow-on exploration. This mapping effort would ultimately examine 22 topic areas before narrowing to six high-value focus areas. The six focus areas were:

- Availability of green energy resources
- Impact on human health
- Privacy and information integrity
- Water availability
- Contrails
- Waste

Of the six selected areas, "Impact on human health" was chosen as the first topic area for CAS investment and development, marking the formal transition from the broad Zero Impact Aviation concept to what would become Zero Impact Aviation – Human Health (ZIA-HH). ZIA-HH would eventually give rise to the Aerial AId activity.

4.2 ZIA-HH Discovery Process

Throughout fiscal year 2023 (October 2022 - September 2023), CAS conducted a discovery effort comprising design sprints in the ZIA-HH topic area. This "Arc of Opportunity Discovery" was executed by over twenty team members representing a multi-disciplinary collaboration across NASA centers, contractor support, and specialized expertise areas.

4.2.1 Team Composition and Leadership

The ZIA-HH synthesis team was structured around three primary components:

Design Facilitators:

- Eric Brubaker (LaRC) - CAS Synthesis Lead
- Steve Crimaldi (GRC)
- LeeAnn Maryeski (LaRC)

[1] The team acknowledged the name "Zero Impact Aviation" might appear to indicate reducing all impacts of aviation, but the ZIA effort was focused on reducing negative impacts of aviation while enabling positive impacts.

Core Team:

- Barbara Burian (ARC) - Human Factors
- David Fuller (GRC) - Systems Engineering
- Shideh Naderi (AFRC) - Flight Systems
- Emma Pierson (GRC) - Materials and Structures
- Rich Walsh (ARC) - Simulation
- David Wagner (LaRC) - Systems Analysis
- Five NASA interns: Neva Bowers, Trifeena James, Jason Li, Robyn Richmond, Vi Nguyen

Emergency Medical Services (EMS) Mini-Sprint Component:

- David Rinehart (GRC) and Tanner Slagel (LaRC) as design workshop representatives
- Beth Rieken (LaRC), Kathleen Bond (ARC), and Rachel Best (ARC) as specialized design facilitators

4.2.2 Research Process and Stakeholder Engagement

The synthesis process involved extensive stakeholder engagement and research activities that demonstrated the agile nature of the CAS discovery methodology. Throughout this process, challenges in the healthcare system as well as opportunities for NASA's expertise and technology to solve such challenges were iteratively refined and reworked through workshops, stakeholder interviews, co-design sessions, and surveys with citizens. Through this process there were over 50 stakeholder interviews, 3 site visits, 9 co-design sessions, 5 workshops, and feedback from 130 citizens. The stakeholder engagement spanned multiple sectors and expertise areas, including healthcare providers, emergency medical services personnel, patients in rural communities, aviation industry representatives, regulatory experts, and various NASA subject matter experts. This approach ensured understanding of the problem from technical, operational, regulatory, and end-user perspectives.

This design thinking approach to opportunity development exemplified the CAS Discovery's strength in transforming broad topic areas into actionable research opportunities. Six Opportunity Concept Reports (OCRs) were made from this process, each addressing different aspects of challenges within the healthcare system. The OCRs were:

OCR1: Rural Airports as Hubs of Healthcare and Community - A systems-level approach to reimagining rural general aviation airports as integrated healthcare, transportation, clean energy, and community hubs aligned with Regional Air Mobility visions.

OCR2: Uncrewed Medical Air Support Systems - Combining medical sensing technologies, locating technologies, and unmanned aircraft systems to improve emergency medical response and increase access to preventative and specialized care in rural and underserved areas.

OCR3: Radiation Shielding - Enhancing shielding technologies and portability of medical imaging equipment to improve accessibility for patients requiring diagnostic imaging.

OCR4: Integrated Patient Transfer and Transportation System - Connecting aviation and healthcare systems to facilitate more efficient inter- and intrahospital patient triage, transfer, and transport.

OCR5: Modular Heavy Lift Vehicle & Mobile Clinic Module - Addressing mobile healthcare provider challenges through development of new modular heavy lift Vertical Takeoff and Landing (VTOL) vehicle capabilities.

OCR6: Aviation-Enabled & AI/ML-Enhanced Organ Delivery - Linking organ matching algorithms with transportation systems to optimize the time-sensitive challenge of organ allocation and logistics.

4.2.3 A CAS Roundtable for Six OCRs

A diverse and expert roundtable panel was assembled to provide evaluation and feedback about each of the six OCRs across multiple domains. Participants of the roundtable included NASA ARMD leadership, external representatives from the National Institute of Health (NIH), National Science Foundation (NSF), and the Federal Aviation Administration, NASA subject matter experts and researchers, healthcare experts, the CAS project team, and the ZIA-HH design team.

The roundtable process was structured to foster collaborative discussion, as well as to gather systematic evaluation of each OCR. Each OCR session followed a structured agenda:

- Problem introduction and context setting
- Individual OCR presentations by synthesis team members
- Collaborative discussions for each OCR
- Cross-OCR comparative discussions
- Systematic evaluation capture

The evaluation employed a dual approach combining qualitative narrative feedback with quantitative Likert-scale ratings across six categories:

1. **Compelling Nature of the Opportunity:** Assessment of the problem significance and opportunity attractiveness
2. **Feasibility:** Technical veracity, plausibility, and achievability within reasonable timeframes
3. **Viability:** Ability to implement and scale, the market potential and fit, regulatory pathway considerations

4. **Desirability:** Compelling societal and stakeholder needs, alignment with beneficiary priorities

5. **Transformational Potential:** Capability to create significant positive change or enable new paradigms

6. **NASA Alignment:** Consistency with NASA mission, capabilities, and strategic priorities.

Additionally, qualitative responses to the topic recieved throughout the roundtable and in the feedback survey were captured.

4.2.4 Selection Decision and Rationale from OCR Roundtable

From the roundtable, OCR1 (Rural Airports as Hubs of Healthcare and Community) and OCR2 (Uncrewed Medical Air Support Systems) emerged as the leading candidates through the combined evaluation approach. While their numerical ratings were not as high as OCR3 (Radiation Shielding), the qualitative discussions and technical assessments identified them as having the strongest combination of societal impact potential, technical feasibility, and alignment with NASA capabilities.

The CAS Project leadership made the decision to advance OCR1 and OCR2 to the "X1" execution planning phase that would build a team and create a two-year effort towards enabling a concept critical to the OCR. Although this selection was based on several key factors, we focus below on the the decision to advance OCR2, since that is the OCR relevant to Aerial AId.

OCR2 was recognized as highly desirable for society. Success, however, would depend on implementation costs, revenue source, financial liabilities, security, and privacy concerns. Additionally, while the compelling use cases in the OCR were strongly aligned with societal needs, some of the proposed solutions are already being pursued. Because of this, the X1 team was instructed to carefully map and ensure distinctiveness from the relevant existing work in industry, other government agencies, and across NASA.

In terms of technical feasibility, reliability and safety of the system were identified as key, but it was not yet clear if such reliability would be feasible given the challenges associated with rapid response and navigation within urban environments, particularly in emergency medical situations and in close proximity to people. Consequently, the effort focused on pushing the capability envelope and proving safety and reliability to make the uncrewed medical air support system transformational. It was also unclear if it was feasible to reliably determine the level of emergency at an EMS scene through a drone, even one equipped with a camera and other instruments.

It was noted that this concept has the issue of regulatory viability, since UAS regulatory barriers are not only formidable in the U.S., but are also critical for unlocking scalability for the medical UAS industry. The consensus was that NASA might have a role to play with relevant agency authorities including the FAA; however, these regulatory barriers would not be the right fit for a CAS execution effort.

Transition pathways for technology developed related to OCR2 were identified within NASA ARMD, including the Advanced Capabilities for Emergency Response Operations (ACERO) project, the System-wide Safety (SWS) project, the Airspace Operations and Safety Program (AOSP), and the Air Traffic Management – eXploration (ATM-X) project. Beyond NASA, the transition potential was identified to be with emergency departments, search and rescue organizations, and the Federal Emergency Management Agency (FEMA). The use case of using UAS for emergency response was identified as something that would strengthen the case for UAM.

The roundtable process also established clear pathways for the non-selected OCRs. OCRs 3-6 were placed in the CAS "backlog/compost" for potential future consideration, with plans for continued analytical work to identify cross-cutting elements and technical challenges that might inform future opportunity development.

4.3 Uncrewed Medical Air Support Systems Exploratory Phase

The Exploratory Phase (X1) for OCR2 "Uncrewed Medical Air Support Systems" began in September 2023, marking a transition from concept development to focused technical and market validation. This phase was designed to refine the broad opportunity concept into a more targeted, executable research activity through intensive stakeholder engagement, technical feasibility assessment, and teaming.

The exploratory phase was structured around two parallel work streams designed to address different aspects of the opportunity maturation. The first was a series of events focused on convening the medical UAS industry. This was a stakeholder engagement and mapping effort to further synthesize concepts in OCR2 that represented a unique NASA opportunity to enable the growth of the medical UAS industry. The second was to form a rapid expeditionary team for a two-year execution activity that would be proposed and detailed in an execution agreement.

4.3.1 Stakeholder Co-Design Process

Building on the initial synthesis work, the X1 team conducted a stakeholder engagement process to find what concepts NASA could move forward to enable the growth of the medical UAS industry and ultimately result in an increased quality of care that people at emergency scenes receive. This effort included 24 co-design sessions with relevant stakeholder groups that included 30 unique external stakeholders. This stakeholder involvement spanned multiple expertise domains, including: 10 industry players (medical UAS companies, technology providers), 7 EMS/dispatch providers (emergency medical services professionals), 4 NASA subject matter experts (technical and domain specialists), 3 fire chiefs, 3 regulation experts (UAS regulatory and policy specialists), 2 emergency management/risk experts (disaster response professionals), and 1 search and rescue expert (specialized emergency operations), and 1 emergency room doctor.

Co-design sessions were often 1-hour discussions around topics and issues related to OCR2. These conversations were structured around strategic research questions designed to identify NASA's unique value proposition and clarify market positioning:

- What constitutes true "whitespace" (meaning it does not represent a redundant effort with existing industry work)?
- Where is industry blocked in current development efforts?
- Who are potential strategic partners for future collaboration?
- What is the desired future state we want to enable?
- What should NASA's specific role be in this ecosystem?

As the co-design sessions progressed, various ideas emerged that presented as viable CAS efforts.

4.3.2 Medical UAS Industry Workshops

After the series of co-design sessions, two consecutive workshops of NASA civil servants were convened. The first was focused on synthesizing all the data from the co-design sessions, and the second was to take the data and the potential concepts for execution and decide which one would be developed into a two-year plan. Figures 2 and 3 show the two concepts that emerged as the best fit for NASA.

"Assured perception" was a concept focused on enabling and assuring computer vision from the perspective of a UAS, towards autonomous scene perception and assessment. Stakeholders expressed interest in such technology but frequently cited its lack of trustworthiness as a barrier to adoption. In addition to this, there were concerns about computer vision for an aerial use case because classifiers were not performant from an aerial perspective due to a lack of data focused from an aerial perspective for emergency medical scenes.

The second concept was focused on utilizing Alternative Position Navigation and Timing (APNT) to gather very precise location information for a UAS so it can operate in areas where quality of traditional position, navigation, and timing signals (like GPS) are degraded or unavailable. From stakeholder discussions, it was clear that in order to have any future concept of widespread UAS usage, APNT must be available so that the position information of a UAS is available and accurate. Stakeholders were interested in the concept, and it was identified as a concept NASA should pursue, but it was ultimately not chosen as the focus for the execution effort. Due to stakeholder feedback and discussion during the workshop, where it appeared that the APNT space for enabling UAS was being pursued by industry, and that it also represented a more traditional ARMD activity, this concept was not chosen. The concept focused on assuring perception was more focused on a transformational visual of future operations and aligned with CAS in its scope as well as its complexity and difficultly, which CAS called a "wicked" problem.

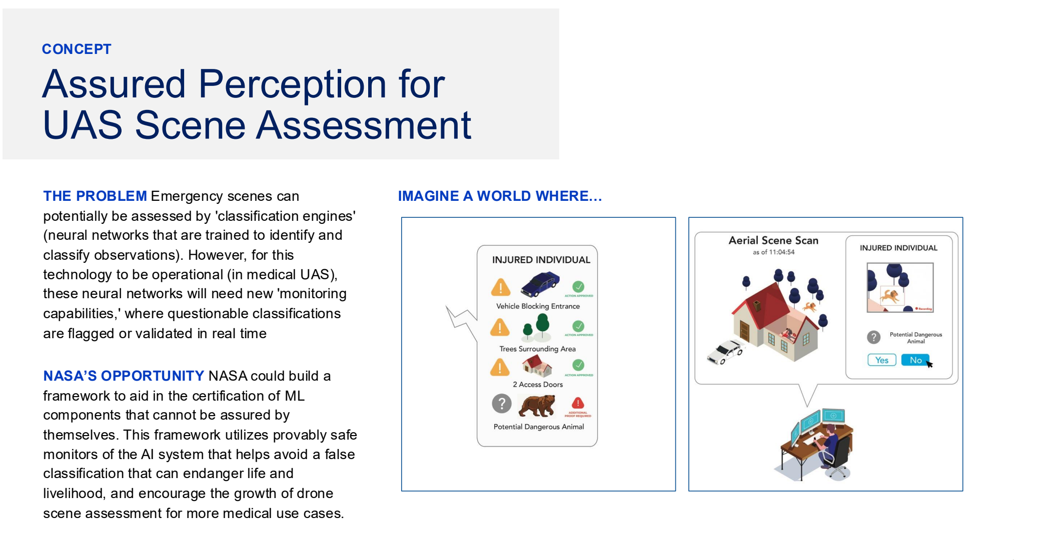

Figure 2: Slides developed during the exploratory phase of Aerial AId for stakeholder co-design sessions

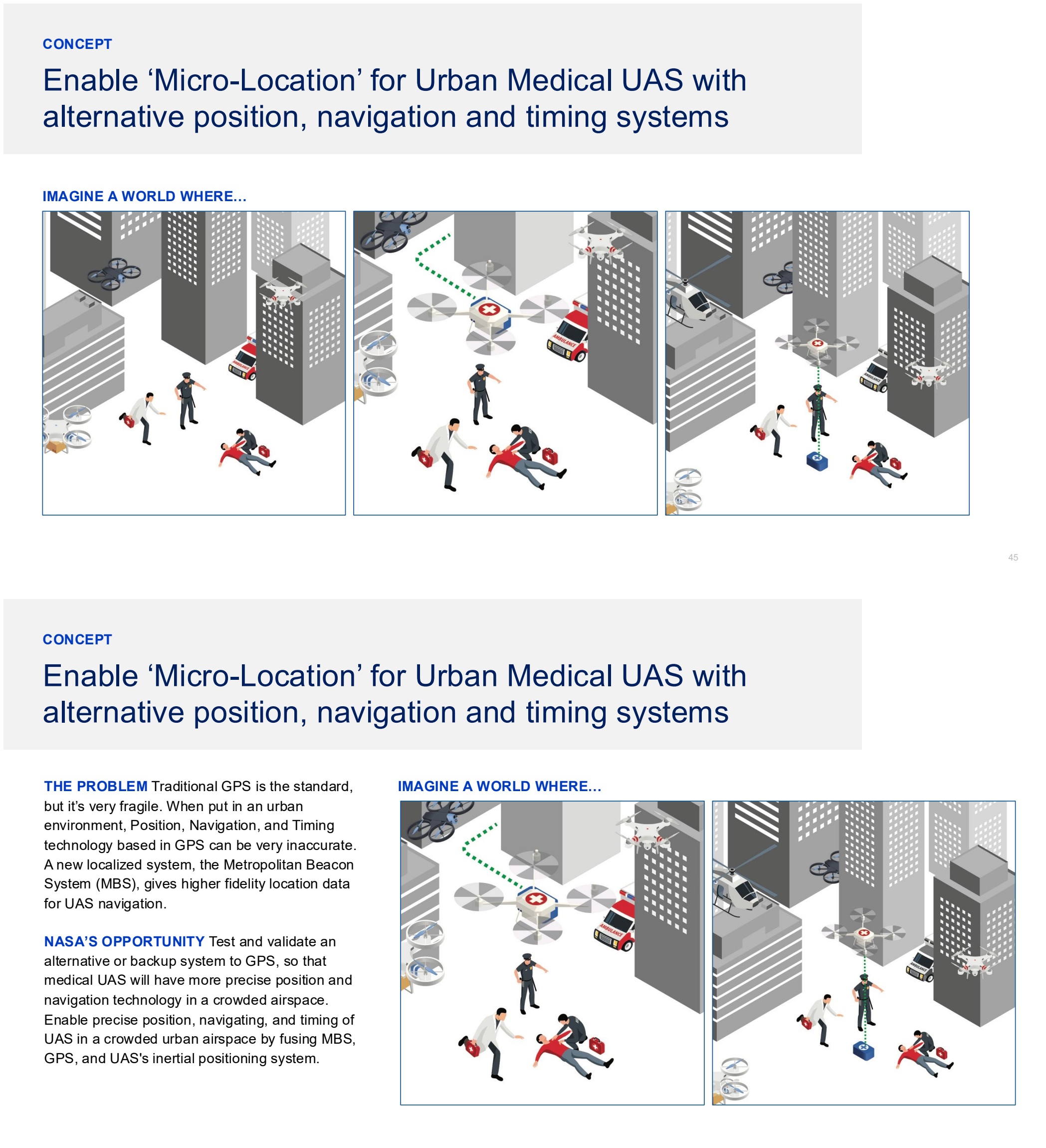

Figure 3: Slides developed during the exploratory phase of Aerial AId focusing on APNT for stakeholder co-design sessions

4.3.3 Assuring Computer Vision and Perception Engines

Throughout the stakeholder engagement process, a clear technical focus emerged around computer vision and automated perception capabilities for emergency medical response. The co-design sessions revealed several critical insights that were categorized according to desirability, viability and feasibility:

Desirability Barriers:

- First responders need higher situational awareness of emergency medical scenes before arriving

- Citizens want privacy protection relative to UAS operations
- Emergency medical services require simple implementation without high integration complexity
- The emergency medical system faces severe resource constraints and personnel limitations
- Medical UAS stakeholders need performant classifiers optimized for aerial perspectives

Viability Barriers:

- Lack of aerial data for training perception engines and insufficient overhead required to generate appropriate training datasets
- Inability to assure AI/ML components for certification and industrial adoption
- Emergency medical industry resource constraints requiring technology to reduce cognitive load
- BVLOS operations as essential prerequisite for unlocking medical UAS industry potential

Feasibility Barriers:

- Difficulty in assuring and certifying AI/ML software components for safety-critical applications
- Uncertainty about machine learning classification effectiveness from aerial perspectives
- Requirements for edge computing implementation that doesn't affect UAS airworthiness, while allowing low-latency integration with navigation and control of a UAS
- Need for automatic environmental perception and mapping without pilot involvement

4.3.4 Market Research Report

The market research conducted for the Aerial AId project revealed significant insights into the current state and future potential of UAS technology for emergency medical response [7]. The overall UAS market is robust and expanding rapidly [8], with a notable emerging sector for DFR applications. Despite this growth trajectory, the research found that current DFR applications are predominantly constrained to law enforcement uses, primarily providing remote video surveillance capabilities [5, 9]. Notably absent was evidence of UAS being utilized by medical first responders for scene assessment purposes, with related medical emergency applications limited to pilot programs for the delivery of medical supplies and equipment [10,11]. This identified gap in the market represented a clear opportunity for the Aerial AId project to address unmet needs in emergency medical response.

While the concept of fully autonomous DFR for medical applications, including UAV ambulances and patient transport drones, has been discussed in the literature, these applications remain largely unrealized in practice. The market research determined that the technological foundation for full autonomy is likely in place; however, several critical barriers have impeded further development and implementation. These include regulatory constraints, safety and privacy concerns, financial limitations, public perception issues, and integration challenges with existing emergency response systems. Despite these obstacles, recent FAA regulatory updates passed in 2023 may accelerate adoption by enabling BVLOS operations [12], a critical capability for effective DFR deployments. Additionally, there are a number of companies with interest in enabling UAS for emergency response such as Dragonfly, Paladin Drones, BRINC, PELA Systems, AI Robotic Drone Solutions, and the not-for-profit MITRE Corporation. Consulting organizations like Skyfire Consulting and UAV Coach offer guides to help public safety organizations establish DFR programs. The non-profit DRONERESPONDERS provides training, certifications, and resources for first responders and holds a number of yearly events and meetings for communities of interest. Government agencies like the Defense Advanced Research Projects Agency (DARPA) and the National Institute of Standards and Technology (NIST) have been involved with UAS for emergency response through hosting various related challenges to spark technological innovation [13, 14].

The market research identified a substantial gap between academic advancements in computer vision and their integration into commercial DFR platforms. Significant progress has been made in developing computer vision algorithms applicable to medical scene assessment, yet there is insufficient evidence of these technologies being effectively implemented in commercial DFR systems. This research-to-market gap is attributed primarily to challenges related to quality assurance for autonomous systems and the lack of application-specific training datasets for AI/ML models [15, 16]. These findings directly informed the Aerial AId project's focus on creating specialized datasets for emergency medical scenarios and developing a runtime assurance framework to ensure AI/ML reliability in safety-critical operations.

To bridge the identified gaps, the market research recommended several strategic pathways: demonstrating research advancements through robust field testing, advocating for improved regulatory frameworks that balance innovation with safety, integrating quality assurance measures for AI/ML components, and engaging stakeholders through effective public communication strategies. These recommendations aligned perfectly with Aerial AId's technical approach, validating the project's focus on dataset development, perception engine optimization, and assurance framework creation as essential contributions to advancing the state of the art in autonomous emergency medical response capabilities.

4.3.5 A Two-year Execution Plan for Aerial AId

From the two workshops and the market research report, a focus on enabling and assuring perception for emergency medical UAS was chosen. From here, an execution planning workshop was carried out to work out the details for a two-year execution effort.[2].

At a high level, the plan was first focused on establishing a partnership and/or access to aerial data related to emergency response operations focused on a particular use case. After data was received, a machine learning dataset could be made to create an aerial classifier. At the same time that a classifier was to be developed, efforts to build a runtime assurance framework that monitored and assessed the trustworthiness of the classifier was to be built. Towards the latter half of the execution effort, a demonstration flight of the assurance framework with the nominal classifier would be carried out on a customized edge hardware device, with results shared with relevant stakeholders. Figure 5 shows a high-level schedule of the execution effort. The execution schedule included a number of CAS's signature "Persevere, Pivot, or Punt" (P3) decision points at 6-month intervals, ensuring regular assessment of progress and strategic alignment throughout the activity lifecycle.

The CAS discovery process was focused on long-term transformation changes. At the same time, end-users with high intensity jobs that are overburdened (as many first responders are), require immediate improvements to their state of operations. With Aerial AId, the research approach sought to strike a balance between near term improvements to DFR, while also focusing on the long-term transformational vision of autonomous DFR operations. Figure 4 from the X1 Medical UAS Workshop in the exploratory effort of Aerial AId in November 2023 graphically shows this approach.

[2]The execution effort was designed to be a two-year effort, the team was told to go into an orderly shutdown after one year, due to the sunsetting of the CAS project in FY25.

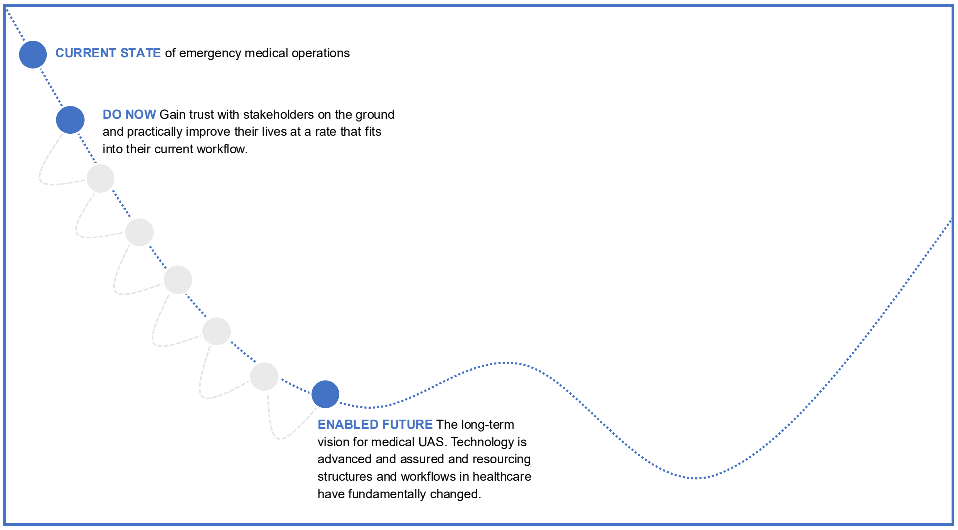

Figure 4: Incremental path towards transformational change in the medical UAS industry from X1 Medical UAS Workshop in the exploratory effort Aerial AId, November 2023

4.4 Transition to Aerial AId Execution

The transition from "Uncrewed Medical Air Support Systems" to "Aerial AId" represented more than just a name change - it reflected a refinement in scope to a specific technical track. The original OCR2 encompassed multiple technical vectors including medical scene assessment, micro-location technologies, and medical package transport. Through the systematic CAS discovery process, these broad concepts converged into a focused technical challenge - assuring and enabling AI/ML perception engines for emergency medical scene assessment from aerial perspectives. This transition to Aerial AId, formalized through a CAS Execution Agreement effective October 1, 2024, represented the culmination of the discovery and execution planning process and the beginning of focused technical development into what would become a one-year execution effort. Figure 5 shows a high level schedule of the proposed Aerial AId execution phase.

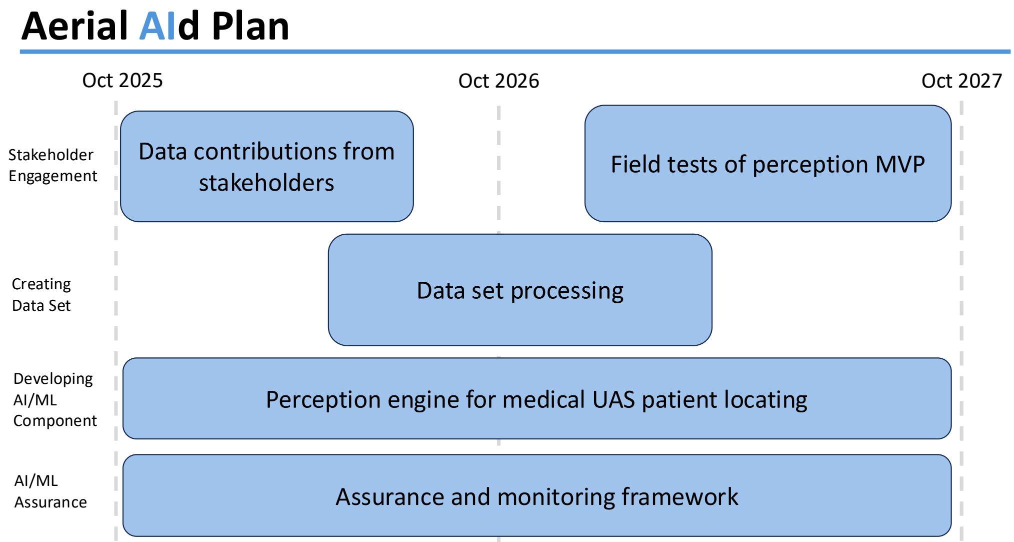

Figure 5: High-level schedule of the Aerial AId execution plan

4.4.1 Multicenter, Multidisciplinary Team Organization

In the standard CAS approach, Aerial AId was structured as a multicenter, multidisciplinary effort drawing scientists and expertise from across NASA's Langley Research Center, Glenn Research Center, Armstrong Flight Research Center, and Ames Research Center.

The team consisted of 20 personnel[3], which included 13 civil servants, 3 contractors, and 4 interns. The team was organized into eight specialized sub-teams reflecting the core technical areas, see Figure 6 below.

[3]Note that most personnel who contributed to the team were not full time team members. Contributions ranged from sporadic advise to full-time team members, with the majority falling somewhere in the middle.

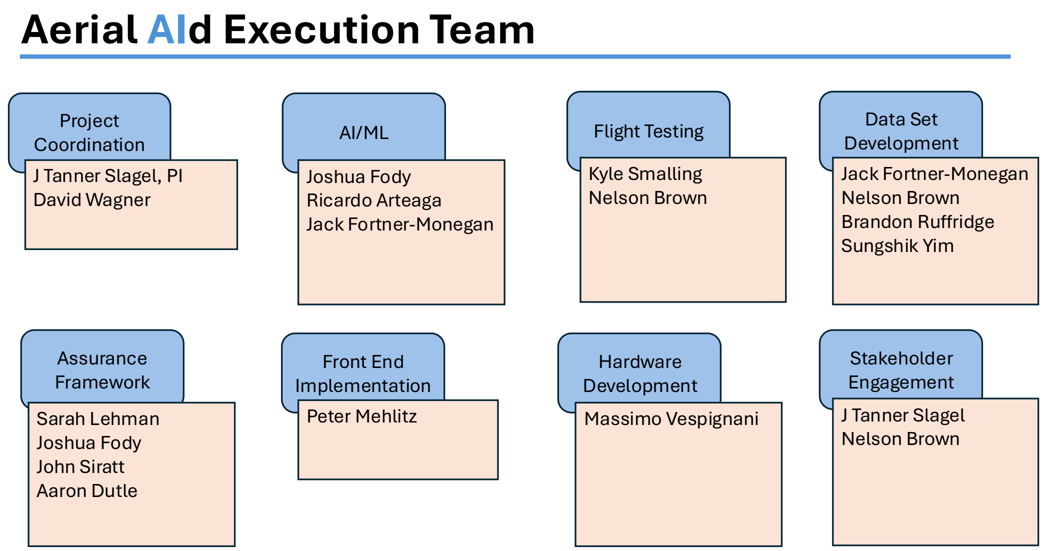

Figure 6: Aerial AId execution team broken down by area of focus

4.4.2 Strategic Alignment and Transition

Aerial AId was aligned with multiple NASA ARMD Strategic Thrusts, demonstrating its integration with broader agency priorities:

- **Strategic Thrust 1:** Safe, Efficient Growth in Global Operations
- **Strategic Thrust 5:** In-Time System-Wide Safety Assurance
- **Strategic Thrust 6:** Assured Autonomy for Aviation Transformation (primary alignment)

While the primary focus of Aerial AId was emergency response, the activity had several cross-cutting technologies that would extend past this space into other general areas of aviation. Particularly, the assuring of AI/ML components, and the subsequent runtime assurance framework could be used for any piece of software that contains such a component.

The System-Wide Safety (SWS) project within ARMD was identified as a primary transition partner for Aerial AId. At the proposed close of Aerial AId at the end of Fiscal Year 2026 (September 2026), SWS was planning to initiate an emergency medical courier safety demonstrator test flight to be completed in 2029. With this in mind, it was the hope to transition Aerial AId technology to aid in this activity. The ACERO project was also identified as a transition partner at NASA, focused on wildfire response operations with UAS use cases that heavily overlapped with emergency medical response operations. Aerial AId planned to make developed datasets available through the ARMD test portal with the possibility for public use, and to make the prototype runtime assurance framework available within ARMD.

Beyond NASA there were a number of UAS companies focused on medical response that would directly benefit from the software/hardware implementation of an assured perception engine as well as the developed aerial datasets; the co-design sessions identified BRINC Drones, the Matador UAS Consortium, Zipline, and Parawave. Note that many firefighters and local police have small UAS teams that could utilize this technology once it is integrated on a UAS. It was noted that the cost and ease of implementing this technology must be kept low for local municipality adoption.

4.4.3 Goal and Deliverable Framework

The overarching goal of Aerial AId was broken down into three deliverables. A schedule showing these deliverables is provided in Figure 7. The three deliverables were:

D1 - the development of a dataset for medical UAS use cases,

D2 - creating an assurance framework for the monitoring of a computer vision component, and

D3 - an integrated software/hardware field test of the assurance framework.

The accomplishment of these three deliverables was broken into sub tasks, and culminated with key event points, described in Figure 8. The completion of these events also marked a time to have a P3 event with CAS, giving the team an opportunity to step back, look at the progress being made, and make a decision to persevere, pivot, or punt future work based on the findings of previously completed tasks.

Deliverable	D	V	F	Task	FY24	FY25				FY26				FY27
					Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1
D1				**Development of data set for medical UAS use cases**										
D1.1				Stakeholder sessions for use-cases										
D1.2				Establish appropriate agreement with stakeholder(s) for data			1							
D1.3				Generation of synthetic data based on use-case(s)			2							
D1.4				Nominal data set for medical UAS use-case(s)					3					
D1.5				Processing for release and archiving							5			
D2				**Software prototype of assured computer vision framework**										
D2.1				Iterative development of assurance component					4		6			
D2.2				Iterative development of computer vision component					4		6			
D2.3				Refinement of integrated assured computer vision framework									7	
D3				**Integrated software/hardware field test**										
D3.1				Hardware specification, acquisition, and assembling										
D3.2				Iterative software/hardware/full stack integration							6			
D3.4				Flight testing and development of edge hardware/software system									7	
D3.5				Stakeholder field tests										8

Exploratory
P3 Event
Closeout

Figure 7: Aerial AId schedule for execution

Event #	Start Date (xQFY2x)	Activity Event (Demo, prototype build, etc.)	Barrier (s) being addressed (i.e., D and/or V and/or F)	Outcome Anticipated
1	2QFY25	Establish Appropriate agreement with DFR Stakeholder(s)	Desirability, viability, feasibility	-Space Act Agreement (or other as appropriate) with a stakeholder
2	2QFY25	Generation of synthetic data based on use-cases	Viability, feasibility	- Creation of relevant synthetic data set based on use-cases of Aerial AId
3	4QFY25	Baseline Aerial Data set for medical UAS use case	Desirability, viability	- Baseline processed data set for Aerial AId use-cases created from raw stakeholder data.
4	4QFY25	Systems integration check-in #1 for perception engine	Desirability, feasibility	-A baseline AI/ML component for patient locating with offline demonstration -Go/no-go decision to proceed on edge-hardware integration -Demonstrate integrated AI/ML component with an incomplete runtime assurance framework -Identify integration obstacles for final D2 MVP
5	2QFY26	Processing, releasing, and archiving dataset for medical UAS use cases	Desirability, viability	-Release processed dataset for Aerial AId use cases to appropriate repository (internal to NASA or release to public) - Processed dataset is anonymized of PII
6	2QFY26	Systems integration check-in #2 for perception engine	Viability, feasibility	- Demonstrate integrated full AI/ML framework, runtime assurance framework, connection with Odin, and user front end using available dataset. -Go/no-go transition to online edge application -Go/no-go on transition to hardware implementation.
7	1QFY26	Final systems integration check-in	Desirability, feasibility	-Flight test scenario and verification/validation of AI/ML assurance framework
8	1QFY27	Complete flight testing and field demonstration	Desirability, viability, feasibility	-Software/Hardware lessons learned from flight testing and field demonstration
CO	1QFY27	Close out	Desirability, viability, feasibility	-Transition to stakeholders

Figure 8: Aerial AId key events

Part II
Dataset Development

This section gives an overview of the dataset development effort within Aerial AId. During this activity, there were two sets of data collection flights as well as work done to generate synthetic data in the Unity graphics engine [17].

The Aerial AId project's dataset development effort represented an approach to address the shortage of training data for aerial human detection and pose estimation in emergency medical response scenarios. Recognizing that existing computer vision datasets such as COCO [18] and KITTI [19] are not well-suited for aerial-based detection applications, the project implemented a data collection strategy focused primarily on human stance detection as a key indicator for emergency medical triage and situational awareness.

It should be noted that other relevant information could be aerially perceived from an emergency medical scene. For example, hazards including structural damage, fire and smoke, and flooding could be valuable information aerially detectable via computer vision. The focus on human stance detection for this activity was chosen to create a single feasible and desirable goal within the project time frame. Additionally, from stakeholder conversations, it was made clear that the most important information a UAS could give a responder about a scene is how many people are present at the scene.

The project's real-world data collection efforts were conducted through two complementary flight test campaigns that generated approximately 500 GB of aerial video data across multiple environmental conditions and emergency scenarios. The first campaign was conducted at NASA Langley Research Center's City Environment Range Testing for Autonomous Integrated Navigation (CERTAIN) [20] facility in March of 2024, while the second campaign was a collaborative effort with the Airborne Instrumentation for Real-world Video of Urban Environments (AIRVUE) activity [21] which was part of the Transformational Tools and Technologies (TTT) Project [22]. The second campaign was completed at Kennedy Space Center in February of 2025. Together, these campaigns produced over 800 minutes of video across 45 different configurations and scene scenarios, resulting in a machine learning dataset focused on human stance detection with approximately $10,000$ images.

In parallel with real-world data collection, Aerial AId developed a synthetic data generation capability to address the inherent limitations of real-world datasets. Such limitations include potential biases related to weather conditions, lighting variations, time of day constraints, and limited participant diversity. The project developed the Synthetic Aerial View (SAiV) tool, a synthetic data generator built upon the PeopleSansPeople High Dynamic Range Image (PSP-HDRI+) [23] framework within the Unity 3D engine. This tool was specifically designed to produce privacy-preserving and fully manipulable aerial imagery focused on human detection and posture estimation tasks for emergency medical scenarios. Additionally, a number of synthetic emergency scenes were generated in the Unity graphics engine as proof of concept and testing videos for future classifiers.

5 Real-World Data Collection Overview

This section describes the two flight campaigns for data collection that were undertaken to capture human stance from an aerial perspective.

5.1 CERTAIN Test Flights

The CERTAIN 1 flight campaign was the first data collection effort for the Aerial AId project, conducted at NASA Langley Research Center's CERTAIN facility in March of 2024. This flight test series comprised 30 successful test flights totaling 480 minutes of airtime, generating 318 GB of video data focused on human stance detection scenarios. Table 1 gives a summary of the flight campaign.

Table 1: CERTAIN 1 flight campaign data collection summary

<table>
<tr><th>Date</th><th>Flight Status</th><th>Subject Type</th><th>Stance/Scenario</th><th>Site</th></tr>
<tr><td>3/4/24</td><td>3 flights</td><td>ATD</td><td>Sitting</td><td>A</td></tr>
<tr><td>3/7/24</td><td>5 flights</td><td>ATD</td><td>Laying, Sitting</td><td>A</td></tr>
<tr><td>3/12/24</td><td>8 flights</td><td>ATD</td><td>Laying, Laying</td><td>A</td></tr>
<tr><td>3/21/24</td><td>8 flights</td><td>ATD
Human</td><td>Difficult Camping,
Laying, Standing</td><td>A</td></tr>
<tr><td>3/22/24</td><td>6 flights</td><td>ATD
Human</td><td>Bike Crash,
All Stances</td><td>B</td></tr>
<tr><th colspan="5">Flight Path Types</th></tr>
<tr><td colspan="2">Orbital Paths</td><td colspan="3">Concentric circles around subject at prescribed distances/altitudes</td></tr>
<tr><td colspan="2">Fly-by Paths</td><td colspan="3">Straight-line approaches with camera angle variations (0°, -45°, -90°)</td></tr>
<tr><th colspan="5">Scenarios</th></tr>
<tr><td colspan="2">Difficult Camping</td><td>7 flights at Site A</td><td colspan="2">Human and ATD amid wrecked camping equipment</td></tr>
<tr><td colspan="2">Bike Crash</td><td>5 flights at Site B</td><td colspan="2">Two bicycles, human, and ATD in various accident configurations</td></tr>
<tr><td colspan="2">Stance-only</td><td>18 flights</td><td colspan="2">Sitting, Standing, Laying</td></tr>
<tr><th colspan="5">Campaign Summary</th></tr>
<tr><td colspan="2">Total Flights</td><td>30 successful</td><td>Total Flight Time</td><td>480 minutes (8 hours)</td></tr>
<tr><td colspan="2">ATD Flights</td><td>16 flights</td><td>Human Flights</td><td>14 flights</td></tr>
<tr><td colspan="2">Test Sites</td><td colspan="3">Site A (grassy field), Site B (concrete walkway)</td></tr>
</table>

5.1.1 Choice of Hardware

All data collection flights utilized a Skydio 2 UAS supplied by NASA Langley's UAS Operations Office (UASOO). This hardware was chosen for its ease of deployment. Since no specialized software was being used, the Skydio 2 was able to be used with no specialized configurations. The Skydio was equipped with a 4K electro-optical camera capturing 30 frames per second. The aircraft's built-in safety features, including automatic return-to-launch functionality upon communication link loss and robust flight control systems made it a safe choice for operations involving human participants. The Skydio 2's maximum flight time of approximately 23 minutes per battery cycle made each flight between 16 and 22 minutes.

5.1.2 Participants

During the beginning of the flight test campaign, human participation was not authorized due to potential hazards in the data collection process. During this time data was collected using an Anthropomorphic Testing Device (ATD) colloquially known as a "crash test dummy." By dressing up the ATD in different outfits and putting it in different stances, a variety of data was able to be collected. However, this approach did have some notable drawbacks: the ATD could not move on its own accord and was therefore rigid through each individual flight path, the ATD was also a fixed weight (150 lb), height (5 ft 8 in), sex (male), and race (Caucasian). While the ATD enabled immediate data collection, its mechanical limitations prevented standing positions. All of these factors could be limiting when trying to develop a robust classifier. To introduce some variability and prevent overfitting to specific visual characteristics, the ATD's attire was systematically changed between different scenario configurations, including variations such as plaid blue button-up shirts with blue jeans, yellow baggy sweatshirts with hooded configurations, and brown sweatshirts with black curly-hair wigs, across 16 flights.

The final 14 flights incorporated four human participants who were selected after human participant activity was approved for this flight test campaign.

5.1.3 Scene Scenarios

Participants were placed in two distinct test sites within the CERTAIN I range: Site A, which positioned participants in a grassy field environment, and Site B, which featured participants on a concrete walking path, providing environmental diversity for more robust algorithm training (see Figure 9).

Figure 9: Operating area for CERTAIN 1 test flights

A total of 18 videos focused on a single human or ATD test subject stance configuration. The stances for the ATD were the sitting, prone, and supine positions. Human participants were able to accomplish these stances, as well as standing and signaling.

The other 12 videos focused on two different emergency scenes to test developed classifiers. The emergency scenarios included "Bike Crash" configurations featuring two bicycles, human participants, and ATDs in various interaction states simulating accident victim assessment, and "Difficult Camping" scenarios depicting participants amid wrecked camping equipment to represent wilderness emergency response situations. These flights included both the ATD and a human participant, see Figures 10 and 11.

(a) “Bike Crash” scenario

(b) “Difficult Camping” scenario

Figure 10: Sample emergency response scenarios

(a) Configuration 1 - front

(b) Configuration 1 - back

(c) Configuration 2

Figure 11: Sample ATD stances

5.1.4 Flight Paths

The flight campaign employed two primary flight paths designed to replicate realistic emergency response UAS deployment scenarios. The fly-by path featured roughly straight-line approaches approximately 50 feet to the left or right of participants at altitudes ranging from 50 to 200 feet, see Figure 12. During fly-by operations, camera angles were systematically varied between 0, -45, and -90 degrees below horizontal rather than continuously tracking subjects, better simulating initial dispatch medical UAS footage where subjects might not be immediately centered in the field of view. The orbital path configuration involved the UAS following a series of concentric circles around the test subject at prescribed distances and altitudes, ranging from 50 to 200 feet in altitude and 50 to 300 feet in horizontal distance, see Figure 13. These two flight paths represented a search of an area for a scene (the fly-by path) and orbiting a known scene (the orbital path configuration).

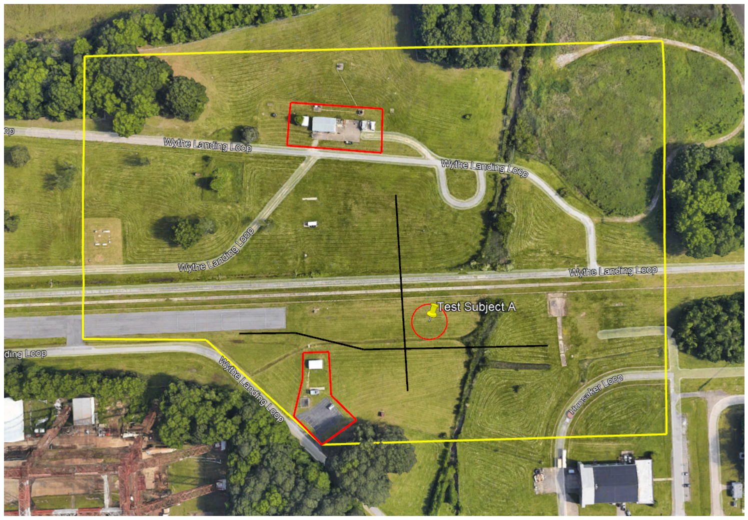

Figure 12: Example fly-by path

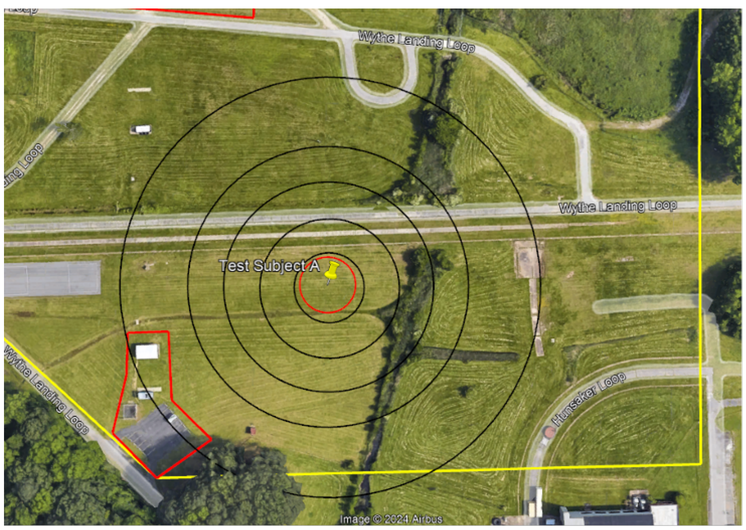

Figure 13: Example circular path

5.1.5 Safety and Risk Mitigation

Hazard mitigation plans were developed and implemented in accordance with NASA UAS Operations Office procedures, addressing identified risks including personal injury, collateral damage, and vehicle fly-away. Flight operations maintained strict altitude parameters between 55 and 300 feet, with no-fly zones established around human participants, buildings, and obstacles such as trees and power lines. The radius changed according to the altitude of the UAS, see Table 2. Additional safety measures included an automatic return-to-launch procedure for link loss exceeding five seconds.

Table 2: Required no-fly zone radius based on UAS altitude

Height of UA (ft)	Required Radius of Keep-Out Zone (ft)
200	40
150	35
100	20
50	10
25	10

5.2 Data Processing from CERTAIN 1 Test Flights

The processing of raw CERTAIN 1 flight data into a usable machine learning dataset began with an automated processing pipeline by the "You Only Look Once" (YOLO) object detection framework [24–27]. YOLO was employed as a first-pass filter to identify frames containing human subjects from the extensive video footage, to reduce the manual annotation burden by automatically detecting when humans were present in the field of view. This approach proved particularly valuable given the scale of the data processing challenges, with over 91,230 frames of interest identified from the raw video footage. The YOLO-based preprocessing served as a coarse-grained method to separate useful frames from non-useful ones, enabling the team to focus human annotation efforts on the most relevant portions of the dataset. However, this automated approach also introduced challenges, particularly when flight test personnel or pilots appeared in the UAS field of view, requiring additional manual review to prevent incorrect labeling of non-test subjects. Additionally, YOLO failed to capture several instances of humans, and so the only initial stance classifications are the images where YOLO has already captured a human, introducing some bias into the dataset.

The first round of data cleaning resulted in a preliminary dataset of 1,598 images, with 1,021 classified as "standing," 347 as "sitting," and 230 as "laying," see Table 3. This processed dataset was then used to train YOLOv8 models, with 1,442 images designated for training and the remainder reserved for validation. This was enough to develop proof-of-concept classifiers, but the need for more labeled data was clear, indicating a need to manually label data in the future.

Table 3: CERTAIN 1 labeled dataset distribution by stance category

Stance Category	Number of Images	Percentage of Dataset
Standing	1,021	63.9%
Sitting	347	21.7%
Laying	230	14.4%
Total Dataset	**1,598**	**100.0%**

5.3 AIRVUE Kennedy Space Center Data Collection

The Kennedy Space Center flight campaign represented a strategic collaboration between the Aerial AId and AIRVUE activities within the CAS and TTT projects under the Transformative Aeronautics Concepts Program (TACP). The flights were conducted in February 2025 and were designed to utilize the capabilities of AIRVUE to advance computer vision datasets for autonomous aviation applications relevant to Aerial AId, specifically human stance classification at emergency medical scenes.

AIRVUE's mission centers on developing computer vision datasets that enable autonomy in aviation applications, encompassing both vehicle-level autonomy through visual navigation and hazard avoidance capabilities, as well as mission-specific applications such as the emergency medical response scenarios central to Aerial AId's objectives. The AIRVUE program has already released multiple datasets via NASA's TechPort that provide valuable resources for innovation in aviation-related computer vision research [28]. Table 4 gives an overview of the data collected accross the flight campaign.

Table 4: Kennedy Space Center AIRVUE flight campaign data summary

Scenario Name	**Camera**	**Duration (sec)**	**Frames**	**Segments**
Grass Field Scenarios				
Separate Position	EO	215	6,450	2
	IR	97	2,910	
Acute Response	EO	150	4,500	2
	IR	89	2,670	
Walking	EO	130 + 16	3,900 + 480	3
	IR	70	2,100	
Separate Position 2	EO	133 + 22	3,990 + 660	3
	IR	54	1,620	
Mass Laydown	EO	75 + 15	2,250 + 450	3
	IR	115	3,450	
Far Group	EO	45 + 81	1,350 + 2,430	2
Beach Scenarios				
Separate Position	EO	158 + 13	4,740 + 390	3
	IR	51	1,530	
Separate Position 2	EO	111	3,330	2
	IR	64	1,920	
Acute Response	EO	177 + 23	5,310 + 690	3
	IR	56	1,680	
Walking	EO	118	3,540	2
	IR	48	1,440	
Separate Position 3	EO	152	4,560	2
	IR	62	1,860	
Mass Laydown	EO	137 + 8	4,110 + 240	3
	IR	54	1,620	
Separate Position 3	EO	145	4,350	2
	IR	47	1,410	
Mass Laydown 2	EO	117	3,510	2
	IR	57	1,710	
Far Group	EO	67 + 11 + 22	2,010 + 330 + 660	3
Campaign Summary				
Total EO Segments	24	2,141 seconds	64,230 frames	
Total IR Segments	13	864 seconds	25,920 frames	
Total Recording Time	37	3,005 seconds	90,150 frames	
Grass Field Coverage	15	1,307 seconds	39,210 frames	
Beach Coverage	22	1,698 seconds	50,940 frames	

5.3.1 AIRVUE Hardware

AIRVUE's hardware capabilities include the custom-developed AIRVUE Development Pod, equipped with wide-angle, forward-facing, and downward-facing cameras designed for scene capture, LiDAR sensors supporting scene mapping, and precision INS-GPS systems providing accurate positioning and orientation data.

5.3.2 Participants

The AIRVUE/Aerial AId test flights utilized 11 total participants, with 9 Kennedy Space Center interns forming the majority of the volunteer pool, supplemented by 2 civil servants. These participants had 5 male participants and 7 female participants of varying heights and body types. All participants were Caucasian, which introduced bias in the resulting dataset. The participants were also skewed towards a younger age range, with 9 of the 11 participants being in their 20s.

5.3.3 Scenario Design

The Kennedy Space Center flight tests were designed to capture baseline human stance data and also capture emergency medical response scenarios across two distinct environmental settings. The scenarios were conducted in both grassy field areas and a beach location.

The scenario design was structured around four primary arrangements. In the first "Baseline" configuration, participants demonstrated individual stance classifications including the standing, walking, laying, sitting, kneeling, and signaling positions. "Walking and Standing" represented a scenario where all participants were walking in small random paths while occasionally sitting down or lying in the prone or supine positions. The "Fallen Person" scenarios simulated bystander response to a single incapacitated individual, capturing the dynamics of emergency scene assessment and initial care provision. The "Mass Laydown" scenarios representing large-scale emergency events where multiple casualties required triage prioritization and limited responder resources had to be allocated effectively. These scenarios were systematically repeated across both the grassy field and beach environments, generating 15 distinct data collection configurations that provided coverage of emergency medical response situations from aerial perspectives. See Figure 14 for a visual of all the arrangements except "Walking and Standing" due to its random nature.

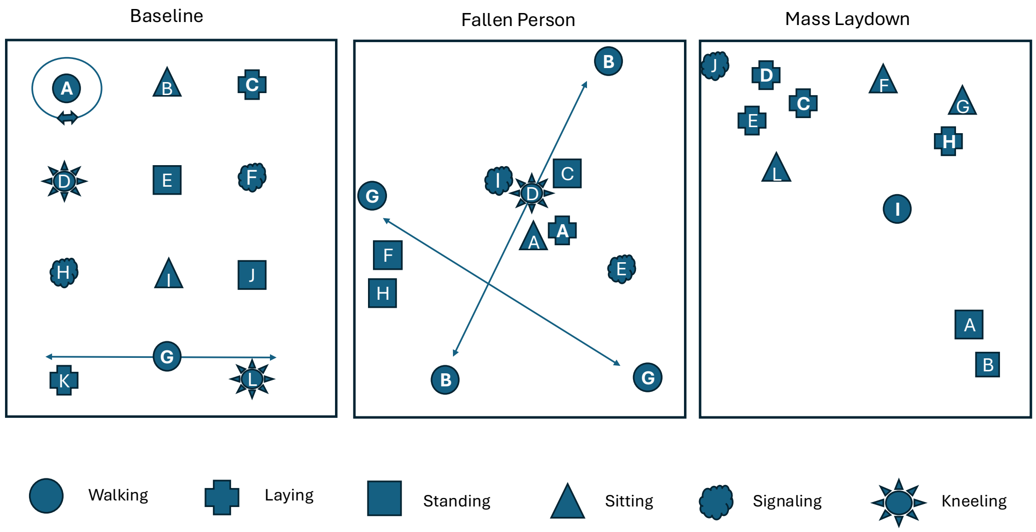

Figure 14: Scenario design for AIRVUE test flights

5.3.4 Flight Paths

Each scenario followed a standardized flight path. The helicopter approach pattern began with a southbound approach at approximately 50 meters above ground level, followed by a direct overhead pass of the staged scenario before transitioning to a controlled hover and 180-degree pedal turn to face south. The aircraft then executed two systematic orbital passes around the scene at a 75-meter horizontal radius—the first orbit dedicated to capturing Electro-optical (EO) color video through the MX-10 camera system [29], and the second orbit focused on infrared (IR) imagery recorded in white-hot mode for thermal signature analysis. The EO segments range from 45 seconds to over three minutes with corresponding IR segments typically lasting one to two minutes. Throughout both orbital passes, the AIRVUE pod's wide-angle cameras maintained continuous recording, while the onboard LiDAR and INS-GPS systems provided precise positioning and orientation data for post-flight analysis. This generated over three hours of multi-spectral recorded data across all scenarios.

5.4 Data Processing from AIRVUE Test Flights

Because of the complexity and multi-participant scenarios, the YOLO-based preprocessing approach was not very successful with AIRVUE data processing. The complexity of multi-person scenarios captured from aerial perspectives, combined with the beach and field environments containing multiple participants in close proximity, resulted in poor performance from the standard YOLO tool. This is expected as it is the motivating aspect of this activity that out-of-the-box computer vision is not very performant from an aerial perspective. The aerial perspective combined with the environmental conditions were challenges that exceeded the capabilities of out-of-the-box YOLO implementations, necessitating a more manual approach towards data annotation.

The project employed Label Studio [30] for manual annotation, revealing the significant labor intensity required for high-quality dataset creation. With 90,150 frames of interest initially identified from the MX-10 video data, the team implemented a sampling strategy, processing every fifth frame to reduce the workload to 18,030 frames. This manual annotation phase was essential in ensuring accurate stance classification, but highlights one of the bottlenecks in creating domain-specific datasets and associated computer vision algorithms. Although some data has been labeled at the time of this writing, a significant amount of raw data that still needs processing remains.

5.4.1 Using NASA Spark Crowdsourcing

Recognizing the limitations of the team towards manually annotating the AIRVUE/Aerial AId flight test data, the project leveraged NASA's internal crowdsourcing capabilities through the NASA Spark initiative [31]. This approach recruited 13 volunteers from across NASA to participate in hand-labeling the multispectral video data using Label Studio annotation tools, distributing the annotation workload across more participants. The crowdsourcing round of data labeling resulted in a preliminary dataset of 5,090 annotated images, and is ongoing.

6 Synthetic Data Generation

The Aerial AId synthetic data generation effort was structured around two objectives to further computer vision development for emergency medical response applications. The first objective focused on emergency scene modeling and demonstration capabilities, enabling the creation of synthetic emergency scenarios that could be used for testing and validating computer vision models in controlled environments. This approach leveraged techniques such as Gaussian Splatting [32] within Unity 3D to generate emergency scenes from previous video data, creating demonstration environments that could showcase capabilities and also be used to prototype and model new technology. The second objective centered on generating synthetic datasets to complement the real-world data collected through the CERTAIN 1 and AIRVUE flight campaigns, with the aim of creating more robust and generalizable computer vision classifiers capable of operating across a broader range of environmental conditions and scenarios than would not be possible if only the real data available was used.

The synthetic data generation approach recognizes the limitations inherent in real-world dataset collection, particularly the susceptibility to biases related to weather conditions, lighting variations, time of day constraints, limited participant diversity, and environmental factors that could compromise the generalizability of trained models. Even with the data collection efforts conducted at CERTAIN 1 and Kennedy Space Center, the resulting computer vision models remained vulnerable to these dataset biases that could limit their operational effectiveness in real-world emergency response scenarios. Synthetic data generation provided a methodology for creating massive, automatically labeled datasets in significantly reduced timeframes compared to real data collection and processing, while enabling full parameterization and randomization of environmental conditions including weather patterns, lighting conditions, time of day variations, object placement configurations, and camera distortion parameters. Additionally, this approach can ensure coverage of scenarios that would be impractical, dangerous, and/or impossible to capture through real-world flight testing, while avoiding the issue of maintaining the privacy-preserving dataset, since participants are not human beings.

6.1 Emergency Scene Modeling and Demonstration Platform

The emergency scene modeling component of Aerial AId's synthetic data generation effort created virtual environments that could serve as testing and demonstration platforms for computer vision algorithms. The Gaussian Splatting approach proved particularly valuable for creating proof-of-concept emergency scenes that maintained photorealistic quality (see Figure 15) while providing the flexibility to modify scene parameters, lighting conditions, and participant configurations as needed for comprehensive algorithm testing.

Figure 15: Example synthetic field environment

With this method, the sidewalk location (Location B) from the CERTAIN 1 test range environment was recreated using the Gaussian Splatting techniques. This synthetic CERTAIN 1 environment allowed researchers to extend the scope of testing beyond the limitations of physical flight operations, enabling the exploration of weather conditions, lighting scenarios, and emergency configurations that would have been impractical or unsafe to conduct during actual flight tests. By maintaining visual consistency with the real test range, the synthetic CERTAIN 1 environment provided a controlled testing platform where algorithm performance could be evaluated under known conditions before deployment in more complex real-world scenarios.

Expanding beyond terrestrial emergency scenarios, the project also developed fully synthetic beach and water-based emergency environments (see Figure 16) in response to insights gained from stakeholder engagement activities, particularly after learning about the New York Police Department's use of UAS technology for shark monitoring and beach safety applications. Committing fully to the aquatic emergency scenarios would require an expansion of the project's scope and the development of a classifier for this application, which was identified for future work. However, it should be noted that DFR applications extend beyond land-based medical emergencies and are well suited for water rescue operations and beach patrol activities.

Figure 16: Example synthetic water environment

6.2 Synthetic Aerial View Framework Development

While each of the previously discussed emergency scene modeling required direct human design, the development of the Synthetic Aerial View (SAiV) framework established a domain-specific emergency medical aerial computer vision application by providing a framework for automatic generation of emergency medical scenes with people making several different stances. SAiV is built on the PeopleSansPeople High Dynamic Range Image (PSP-HDRI+) framework, which itself extends the original PeopleSansPeople (PSP) framework developed in [23] within the Unity engine ecosystem. The PSP and PSP-HDRI+ frameworks serve as the data generators for SAiV, resulting in the production of high-quality synthetic images through parameterized control over lighting and camera settings, while generating ground truth annotations including 2D and 3D bounding boxes, human pose classification, human pose keypoints, and both semantic and instance segmentation masks. However, the SAiV approach extended this to an aerial perspective and with domain randomization techniques, as detailed in [33] to generate random emergency scenes.

6.2.1 Aerial-Specific Adaptations and Constraints

Unlike the synthetic images generated in standard PSP-HDRI+ applications which typically focus on ground-level or eye-level human interactions, SAiV was specifically designed to generate synthetic aerial imagery captured from approximately 50 feet above ground level, reflecting the operational altitude range commonly employed in DFR missions. SAiV constrains the camera orientation angle (α) to fall within a range of 0 to 25 degrees from the vertical downward direction. The camera height parameters were calibrated to operate between 14.5 and 16 meters above ground level, with camera positioning randomized within a 23- to 28-meter radius from the scene center to provide coverage while maintaining the scale and resolution characteristics necessary for effective human detection and pose estimation. Additionally, the field of view was configured to vary between 60 and 90 degrees horizontal, accommodating the optical characteristics of commercial UAS camera systems commonly deployed in emergency response operations. These parameters were derived from analysis of real-world flight test data collected during the CERTAIN 1 and AIRVUE campaigns, ensuring that the synthetic data maintained fidelity to actual operational conditions.

The Unity Perception package, which serves as the foundation for SAiV's data generation capabilities, provides ground truth annotation systems that automatically generate bounding boxes, segmentation masks, and pose keypoint data alongside high-quality red, green, and blue (RGB) imagery, eliminating the labor-intensive manual annotation processes that would otherwise be required for equivalent real-world datasets. Because of this, no hand annotating is required when generating synthetic data in SAiV. Figures 17, 18, and 19 show examples of the SAiV environment without annotations, with bounding box annotations, and with semantic segmentation annotations, respectively.

Figure 17: Example SAiV environment

Figure 18: Example SAiV environment with bounding boxes

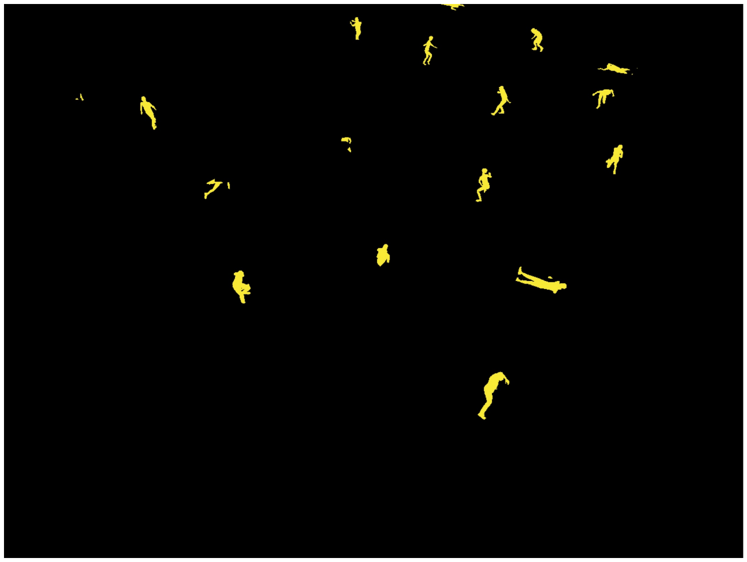

Figure 19: Example SAiV environment with semantic segmentation

6.2.2 Extensions and Connection to the Robotic Operating System

Unity's simulation capabilities enabled SAiV to extend beyond simple RGB image generation to encompass multi-sensor data collection that mirrors real-world UAS sensor suites. The platform's implementation includes RGB cameras, LiDAR sensors (including specific support for Ouster OS0-32 configurations), and radar sensors for future applications, providing a complete sensor simulation environment that accurately reflects the data collection capabilities of modern emergency response UAS platforms. For example, Figure 20 shows Lidar data generated within the SAiV environment in Unity.

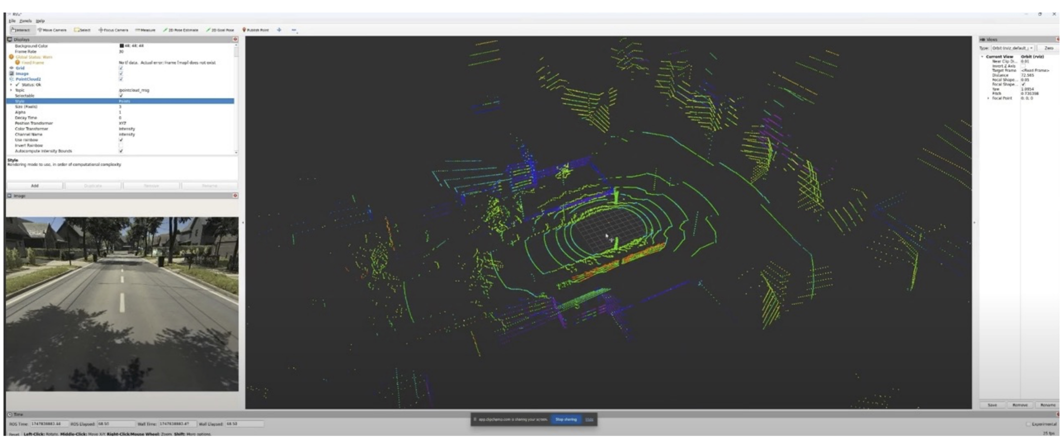

Figure 20: Lidar data generated in SAiV

These simulated sensors publish their data streams to external processing systems, enabling real-time evaluation of computer vision algorithms including YOLO-based detection systems and generative AI models. The multi-sensor simulation capability can prove particularly valuable for testing algorithm performance across different sensing modalities and for validating sensor fusion approaches that combine visual and LiDAR data for enhanced emergency scene assessment, while providing the controlled environment necessary for systematic performance evaluation across varying environmental conditions and sensor configurations.

Additionally, Unity has the ability to connect to the Robotic Operating System (ROS), which includes a testing framework that enables efficient switching between real-world and simulated sensor data for perception algorithm development and validation [34]. This integration supports real-time sensor data processing, allowing perception algorithms developed in the synthetic environment to operate similarly to when deployed on actual UAS platforms equipped with corresponding sensor suites. The ROS integration capability enables research teams to develop, test, and validate complex perception algorithms in controlled synthetic environments before deployment in costly and potentially dangerous real-world flight operations. The framework's ability to replace real-world sensor inputs with synthetic equivalents while maintaining identical software implementation provides opportunities for algorithm assessment, rare scenario testing, and validation across environmental conditions that would be impractical to reproduce through traditional flight-testing methods alone. This approach presents synthetic environments not just as a way to generate data but as development and testing environments.

Part III
Computer Vision for Emergency Response

Computer vision is one of the ways an ML-enabled system can perceive its surroundings. Typically, this involves sampling frames from an RGB or IR camera, passing those frames to an ML model (usually a neural network), and then parsing out the results. If the ML model is doing classification, the model will return a label saying what the principal subject of the image is. If the ML model is doing object detection, the model will get a classification label and a bounding box around each object in the incoming image. If the model is doing semantic segmentation, it will outline each object according to its natural shape, rather than a simplified bounding box. These fundamental computer vision capabilities have evolved significantly in recent years, with a general shift toward computationally complex algorithms based on Convolutional Neural Networks (CNNs) and Deep Learning (DL), which are particularly well-suited to tasks such as object detection and tracking in challenging operational environments.

These kinds of visual perception tasks can be useful in emergency response scenarios, enabling rapid scene assessment by identifying objects of interest within the scene and passing information back to human operators. For example, a vision model could be used to identify humans within a scene, flag environmental sources of danger such as fires or flood waters, or warn of specific threats such as wild animals or active shooters. By capturing this kind of information, computer vision models can improve situational awareness and preparedness of first responders before they ever arrive on scene. However, the application of computer vision to aerial emergency response presents unique challenges that distinguish it from ground-based applications, including small object resolution due to operational altitudes, an aerial viewing angle uncommon in standard object detection, moving camera platforms that introduce motion blur, and complex backgrounds that can hinder accurate detection and classification.

This section of the report details the performance gaps in out-of-the-box (OOTB) computer vision systems on aerial imagery, provides an overview of computer vision prototypes built during the Aerial AId execution phase, including a framework for human stance detection. Limitations of these models and challenges are mentioned as well.

6.3 Performance Gaps in General-Purpose Computer Vision for Aerial Applications

The fundamental challenge encountered with OOTB YOLO models applied to aerial imagery stems from their training foundation on the Common Objects in Context (COCO) dataset [18], which represents a ground-level perspective of everyday objects and scenarios. Additionally, the COCO dataset does not contain data for emergency medical use cases, lacking categories relevant to human poses, medical equipment, emergency vehicles, and other critical elements essential for DFR operations. It should be noted the COCO dataset has images of humans, so YOLO detects humans in a frame, but this is primarily from the perspective of a human in front of the camera. The COCO dataset does not have objects from an aerial perspective, leading to no classification or misclassifications when the objects are viewed from the 50-400 ft altitude range typical of emergency response UAS operations. Since the model was not trained on aerial imagery, it struggles with the scale relationships, spatial orientations, and contextual cues that characterize aerial imagery.

6.3.1 Missed Detections and Classification Errors

Figures 21 and 22 demonstrate the limitations of OOTB computer vision systems, in this case YOLOv8, when applied to aerial data. The aerial perspective introduces a range of classification errors that highlight the inability of models to adapt to the unique viewing angles encountered in DFR operations. The right image in Figure 21 shows how the model consistently misclassified the metal roof of a pole barn as a "bench," showing how patterns and features that define objects from a ground-level perspective do not translate to accurate detection when viewed from above at a higher altitude. This type of misclassification represents the issue where models attempt to map aerial visual features to the closest matching ground-level object category from their training data, often producing nonsensical results, like the "giraffe" classification in Figure 21. Such algorithmic errors become immediately apparent to human observers, but prove particularly problematic in emergency scenarios where accurate scene assessment remains critical for first responder safety and effective triage operations. Additionally, the more heavily the operations of the system rely on the computer vision with less human involvement, the more disastrous such results can be. For a computer vision system to perform effectively from an aerial perspective, it must undergo training on data captured from aerial angles and relevant to the specific application domain.

Figure 21: OOTB classifier (YOLOv8) trained on non-aerial data attempting to process aerial data

Figure 22: Difficulty of classifying humans from an aerial perspective on non-aerial trained models

6.3.2 Classification Thrashing and Temporal Inconsistency

Beyond static misclassification issues, the OOTB models exhibited a phenomenon known as "classification thrashing," where detected objects rapidly oscillate between different class labels over short time periods, creating unstable and unreliable outputs unsuitable for emergency response applications. This temporal inconsistency happens in part due to camera movement from the UAS flight path in the viewing angle. Figure 23 shows an example where along a straight-line path, the "difficult camping" scene is classified as a "kite" and a "person" several times over a small amount of time. In the case where there are changes in the situation on the ground, the classification changing might be appropriate (Figure 24); however, it must represent classifications that are relevant to the scene. In Figure 25, a standing individual is signaling, and the classification thrashes (more slowly) between "standing" and "laying," because the model has not been trained on the "signaling" stance.

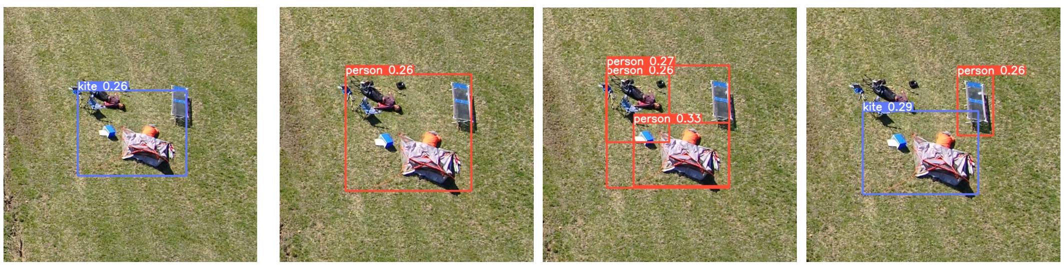

Figure 23: Example of classification thrashing

Figure 24: Correct transition between classes

Figure 25: Classification thrashing due to missing stance (signaling)

6.3.3 Scale-Induced Misclassifications

Even when the model has relevant classifications such as "people," it fails to recognize humans from an aerial perspective due to scale differences. People appear much smaller in aerial imagery than in the ground-level images used for training, resulting in bounding boxes that represent inappropriately large areas. Figure 23 illustrates this phenomenon in the "difficult camping" scenario, where the model generates a "person" classification with a bounding box that would represent an implausibly large human figure given the UAS altitude and viewing angle. The same image demonstrates additional scale-related errors, such as the incorrect "kite" classification that encompasses an area far too large to represent an actual kite at that distance and perspective. These scale mismatches occur because the pre-trained models rely on object size relationships learned from ground-level imagery, where objects fill a much larger portion of the image frame.

7 Overview of Aerial AId Computer Vision Models

Throughout the execution activity of Aerial AId, different models were trained to be prototypes for the runtime assurance framework. Although none of the models have been fine-tuned or trained on sufficient data to be robust enough for deployment, they served as proof-of-concept artifacts, prototypes to test the runtime assurance framework, and nominal models for more advanced versions to be compared to. It should be noted that the goal of Aerial AId was not to make computer vision models ready for deployment; this task is better suited for industrial partners with the time and computation power to make a robust model. The goal of Aerial AId was to make technology to assure such computer vision components, so reference and baseline models were needed on which to test the framework.

The Aerial AId team primarily used YOLO versions 8 and 11 to create the prototype computer vision models. This decision was for several reasons. First, YOLO is quite easy to implement off the shelf, promoting an ease of transferability to applicable research projects. Second, YOLO has different architecture sizes: nano, small, medium, large, and extra large, that allowed the team to find a model that was lightweight enough to run on an edge device. Third, it is easy to perform transfer learning, i.e., to retrain selected pre-trained YOLO weights on new data, offering a way to adapt a YOLO model for aerial object classification without the computational burden of training from scratch. In addition to object detection and classification, YOLO has semantic segmentation, object tracking, and pose estimation, each of which were used during the development of an aerial computer vision system.

7.1 Human Locating and Stance Identification

The primary use case for computer vision explored in the execution phase was focused on human stance identification. This was motivated by conversations with first responders that made it clear the most important information at a scene was the number of patients. At a high level, this might seem like the number of humans; however, it is often the case at a scene that there are bystanders who do not need care. Not only is human identification important, but the designation of care they need is also paramount to the resources allocated to the scene. The most refined version of this technology would mimic triage, which is a well-defined decision procedure (first responders call it the triage algorithm); however, it is unlikely that full triage is possible from just EO and IR camera frames. What can be extracted is stance information - which might be fed into some initial triage decision making. For example, if a person is ambulatory, they are tagged at a green triage level, which can be done from an aerial computer vision model. The human stance model we developed was focused on different stances: standing, sitting, prone, supine, signaling, kneeling, and walking (see Figure 26). In addition to a single computer vision component, we used a multi-model approach that also utilized keypoint information, which extracts specific locations on an individual's body such as joints and hands.

Figure 26: Baseline human stances

7.2 Structural Damage and Obstructions

Another use case explored was one identifying structural damage of buildings and obstructions primarily around roadways for disaster response activities (see Figure 27). This use case was motivated by data received from the Boone County Fire Protection District (BCFPD) and Missouri Task Force 1 (MO-TF1) Disaster Situational Assessment and Reconnaissance (DSAR) team, from the St. Louis tornado aftermath in May 2025. Receiving fast information about structural damage gives first responders locations that must be surveyed and potentially evacuated, and obstructions in the road give them information about viable paths to these structures.

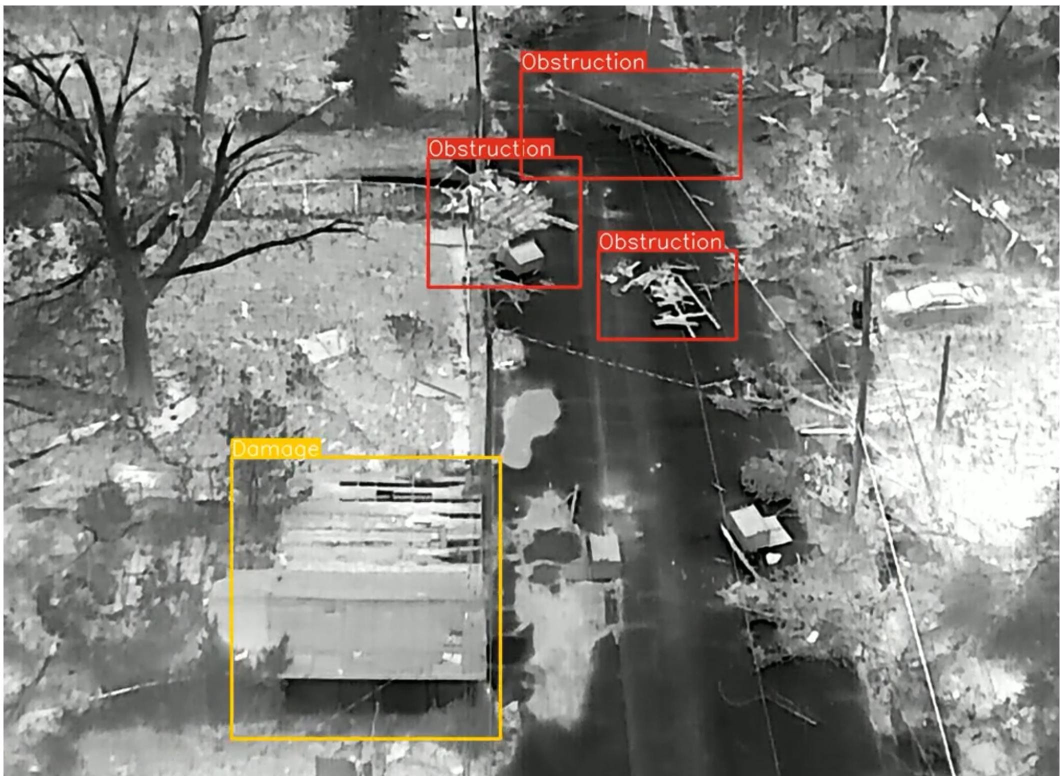

Figure 27: Prototype damage and obstruction classifier[4]

7.3 Car Crash Detection

Car crashes happen every day all over the U.S.; the faster the response time to a critical crash, the higher chance of positive outcomes for the patients at the scene [35]. Rapid identification of car crashes was prototyped over the summer of 2025, and Figure 28 shows the results of a proof-of-concept classifier developed during that time. Instead of UAS data, the team explored data from free traffic cameras offered by the Virginia Department of Transportation [36]. This represented a small

[4]Original image source: BCFPD/MO-TF1 DSAR team, https://dsar.motf1.com/

deviation from the main focus of computer vision on UAS for Aerial AId, but was recognized as having a high potential for use.

Figure 28: Proof-of-concept car crash detection[5]

8 Human Stance Classification Development

The primary goal of Aerial AId's computer vision system was to provide a component that could assist emergency medical responders in rapidly assessing and prioritizing individuals at an emergency scene. However, it became apparent early in the development process that comprehensive medical triage cannot be entirely accomplished through computer vision alone. Traditional medical triage protocols require assessment of things like respiratory rate, pulse quality, and other bodily functions that cannot be reliably determined from aerial imagery [37]. It should be noted that with advanced sensors, perhaps more complete triage could be performed, which is being explored [14].

The framework's design incorporated an additional "towards triage" assessment capability that would be applied at the final stage of the classification system, translating the detected human stances into simplified emergency response categories. Individuals observed walking could be automatically assigned a "green" triage level, indicating minimal immediate medical intervention requirements and potential availability to assist in emergency response activities or provide information about the incident. The system was designed to consolidate the various stance classifications into three primary categories for practical emergency response applications: "upright" (encompassing standing, walking, and signaling poses), "on-the-ground" (including sitting and kneeling positions), and "down/immobile" (covering laying, prone, and supine positions). These three simplified classes would be displayed as green, yellow, and red bounding boxes, respectively, in visual results.

8.1 Two-Stage Detection and Classification Pipeline

In order to extract human stance information, Aerial AId adopted a two-stage approach that could more reliably detect humans in aerial imagery and accurately classify their stances for emergency medical assessment than that of a single pass of object classifier alone, see Figure 29. The first stage employed a transfer-learned YOLOv8 model as the initial category classifier. This initial classifier goal was to aerially detect the presence of humans at the scene and provide bounding box localization around each detected individual, while simultaneously generating an initial stance classification estimation based on the overall visual characteristics of the detected person. The process to train this model is outlined in Figure 30. The development of this classifier predated the ML dataset created by the CERTAIN 1 and AIRVUE test flights, so it was initially trained on the NOMAD dataset. After this, it would be trained again on the CERTAIN 1 and AIRVUE data.

[5] Original image source: Traffic Cam Watch on YouTube/Safetyvid, https://www.youtube.com/@TrafficCamWatch

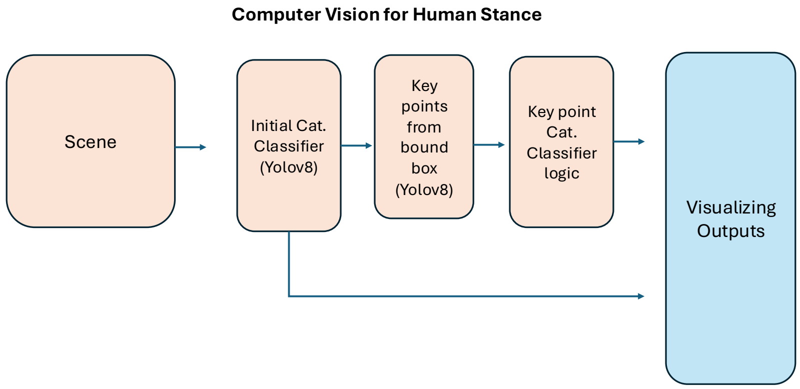

Figure 29: Two-Stage detection and classification pipeline for human stance

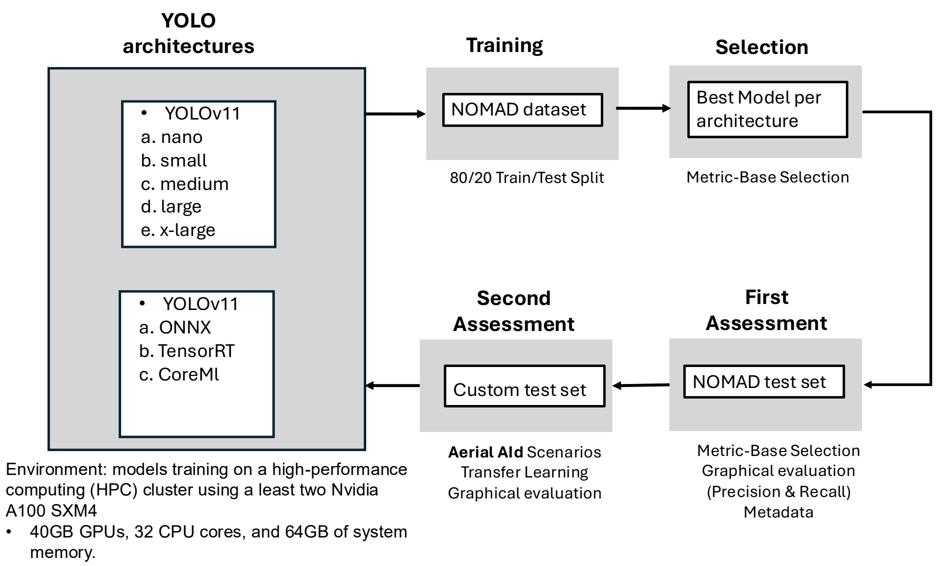

Figure 30: Plan for development of human stance classifier

The second stage of the pipeline focused on extracting human pose keypoints within the bounding boxes identified by the initial classifier for further stance evaluation. This also utilized YOLOv8, but instead of the object detection and classifier model, the bounding box was used as input for the model and the keypoint identifier was used. Using the bounding box coordinates provided by the first stage, the keypoint detection system of YOLOv8 detects key points from a detected human subject, including head, shoulder, elbow, wrist, hip, knee, and ankle positions.

The goal was to process the extracted keypoints through a custom keypoint logic system that would analyze anatomical landmark relationships to determine human stance categories. At the time of this report, however, this logic system remained unfinished with limited testing.

8.1.1 NOMAD Dataset Implementation and Validation

The two-stage computer vision framework was trained and tested using data from Natural, Occluded, Multi-scale Aerial Dataset (NOMAD), a benchmark dataset specifically developed for emergency response scenarios with emphasis on search and rescue operations. The goal was to use this data with the Aerial AId datasets to make a robust classifier. NOMAD represents one of the most extensive publicly available aerial human detection datasets, consisting of 42,825 frames extracted from high-resolution 5.4K videos featuring actors performing sequences of hiding, laying, and walking under different levels of occlusion by obstacles. The dataset's design includes videos collected at 30 FPS with 5,472 x 3,078 pixel resolution from five progressively more distant locations ranging from 10 meters to 90 meters, providing realistic scale variations that closely matched the operational parameters of DFR applications. Importantly, NOMAD featured demographic diversity with 100 actors maintaining a 50/50 male/female distribution aged 18 to 78 from various racial backgrounds, filmed across twelve locations, providing variety in natural and man-made environments including fields, lakes, forests, and schools under different weather conditions.

Figure 31 shows a proof-of-concept implementation of this two-stage classifier. On the top we see a classifier, based on the NOMAD classifications, that identifies a walking person; the key point extraction and logic deduces they are standing, which is in line with the original class. The second image image shows a mismatch between the initial classifier and the keypoint extraction logic classifier. The person is said to be walking, but the keypoints indicate the individual is laying down.

Figure 31: Proof-of-concept two-stage stance framework[6]

9 Limitations of Aerial AId Models and Generalization Challenges

While the transfer learning results demonstrated the use of computer vision for DFR operations, it is important to acknowledge the limitations of this work and the future work necessary to make such a model robust and trustworthy. The enhanced performance observed when testing the transfer-learned models primarily represented an exercise in interpolation rather than extrapolation, as the model was trained on data collected during the same flight campaign, using the same UAS platform, at the same location, and under similar environmental conditions as the test data. Despite employing proper train/test/validation split methodologies to

[6]Original image source: NOMAD dataset, https://github.com/ArtRuss/NOMAD

prevent data leakage and ensure statistical validity, the inherent similarity between training and testing datasets meant that the model had been exposed to the specific visual characteristics, lighting conditions, terrain features, and camera perspectives present in the test scenarios. Additionally, even with the aerial data that was collected and turned into a ML dataset (much of which wasn't available until the end of this effort), the amount of data needed to make a classifier robust enough for deployment is much greater. This limitation highlighted a fundamental challenge in developing robust computer vision systems for aerial emergency response; while domain-specific training could dramatically improve performance within known operational parameters, the model's ability to generalize to new locations, different UAS platforms, varying environmental conditions, or alternative emergency scenarios remained largely unvalidated. To some degree, a runtime assurance framework can help offset some challenges an AI/ML computer component might experience due to such factors, which will be discussed in Part IV. However, the model needs adequate performance across a wide range of use cases, requiring extensive training data and computational power. The recognition of this limitation became a driving factor in the project's subsequent development of synthetic data generation capabilities, as synthetic datasets offered the potential to expose models to a much broader range of scenarios and environmental conditions than would be practical or possible to capture through real-world flight testing alone.

Part IV
Assuring Aerial Perception

The difficulty providing traditional safety assurances for AI/ML software components represents a barrier to their successful deployment in UAS applications like DFR. This part of the report gives an overview of the efforts towards the assurance of AI/ML components, specifically focused on computer vision, with the Aerial AId activity. The primary thrust in this effort was building a Runtime Assurance (RTA) framework, with the goal of having monitors that focus on model behavior and properties of incoming data to raise a warning when it appears an unsafe or untrustworthy state is going to be reached. The second thrust of this effort was to apply formal methods towards the assurance of AI/ML models. This includes a formal development of the simplex RTA framework and the beginnings of a convolution library in the Prototype Verification System (PVS) theorem prover [38, 39].

10 Runtime Assurance for Computer Vision

10.1 Runtime Assurance Approach

Runtime Assurance (RTA) represents an architectural framework for safety-critical systems that enables the integration of advanced but potentially unreliable components, such as AI/ML algorithms, while maintaining system safety through continuous monitoring and intervention capabilities [40–44]. The fundamental principle of

RTA can be understood based on the simplex architecture [45–51], where a system operates using an advanced controller or component (such as a neural network-based computer vision system) while being continuously monitored by trusted runtime monitors that can detect property violations or anomalous behavior, see Figure 32.

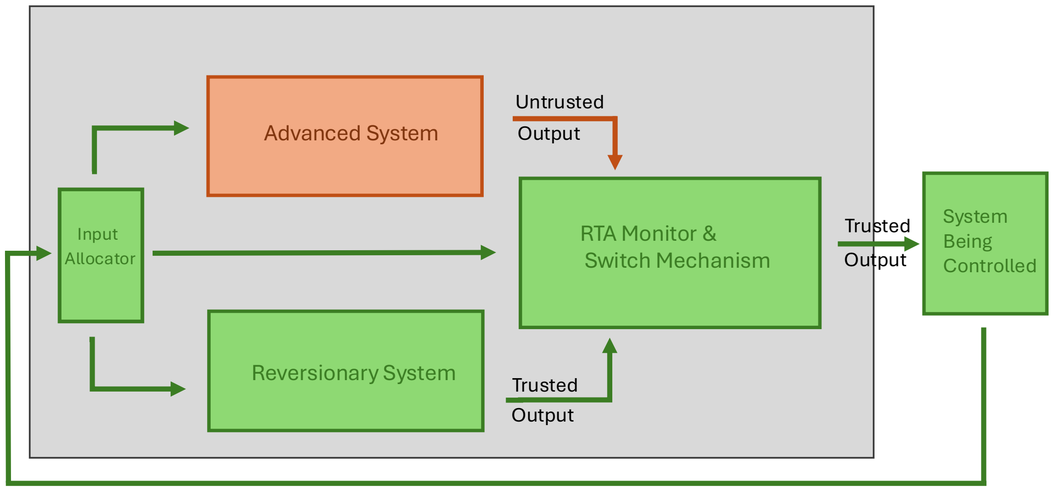

Figure 32: Example of RTA framework, adapted from Schierman et al.'s research [44]

When the monitoring system identifies that the advanced component has violated safety properties or is operating outside acceptable parameters, the RTA framework automatically switches control to a trusted, but potentially limited, reversionary system, ensuring that critical safety requirements are maintained even when the primary system fails or behaves unexpectedly. This approach has gained particular significance with the rising interest in utilizing AI/ML components in aviation applications, leading to published guidelines from both ASTM International and NASA on the use of RTA in systems that rely on AI/ML to operate safety-critical components of aircraft [52, 53].

In reality, the structure of a runtime monitoring system can be much more complex than the simplex method. The proposed RTA framework here is focused on implementing an RTA system that has multiple monitoring approaches to provide more extensive oversight of computer vision performance. When there are multiple monitors, there must be a logic to process the outputs of each to make a decision if the results of the AI/ML component are trustworthy or not and what action to take. This logic must be implemented so that the multi-process implementation does not introduce significant latency penalties.

The key benefits of implementing RTA for DFR operations is being able to have monitors on a high performance, but difficult to assure, computer vision component, where a "flag" is raised that tags untrustworthy results. The RTA approach enables emergency medical UAS systems to leverage the advanced capabilities of state-of-the-art computer vision models, such as human detection and pose classification, while providing safety nets that reduce the risk of failures from happening. This is particularly crucial in emergency medical scenarios where incorrect classification of scene information may cause an inappropriate response, leading to poor patient outcomes. The RTA framework ensures that DFR systems can continue to provide valuable situational awareness even when individual components encounter edge cases or operate outside their trained performance envelopes. In the near term, when the RTA system encounters a potential unsafe or untrustworthy result, DFR operations could loop in an operator for human review. This would make the human reviewer the reversionary controller in the framework. However, as these operations become more autonomous, the reversionary system might be a more performant classifier existing off-board the UAS that can resolve the detected issue at hand.

10.2 System Architecture for Human Stance Detection

Figure 33 shows the RTA framework for the human stance detection prototype described in Part III. In this architecture, there are two separate version of YOLO, and there can be monitors that observe either component for safe behavior. Additionally, a monitor can compare results from both of the components and make a decision based on this aggregate information. One example of this is a mismatch between the initial stance classification of an individual, and the derived keypoint stance of an individual. If these are not the same, that indicates a certain level of uncertainty that might warrant a human review. Table 5 gives an overview of potential monitors for this framework. At the time of this report, a full implementation of this RTA framework for human stance has not been completed, nor have the suggested monitors been implemented. The rest of the section will detail an implementation of an RTA framework applied to a single computer vision component.

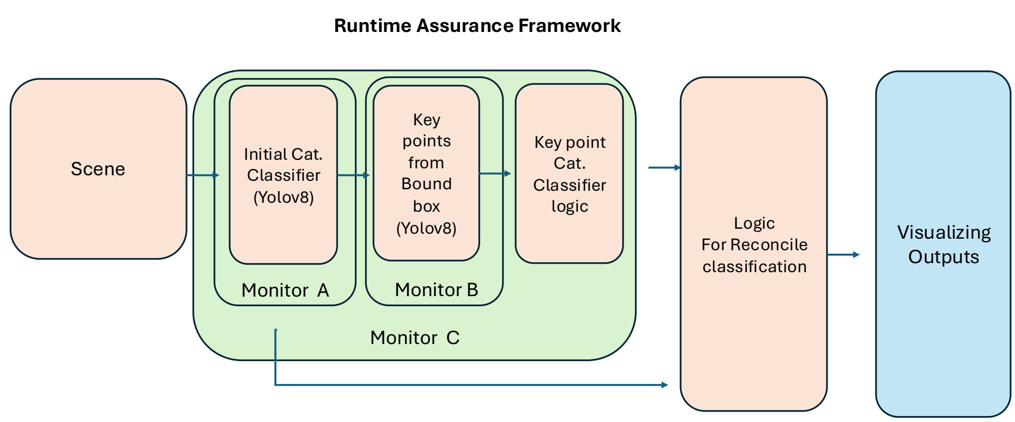

Figure 33: Example of RTA framework applied to two-stage stance classification system

Table 5: RTA monitors for computer vision systems

What is being monitored?	**What is the monitor?**	**Example**
YOLOv8 Initial classifier (bounding box)	Geometric check for bounding feasibility (based on altitude, camera angle, hardware specs, size of bounding box)	A person is detected but the bounding box is 100 pixels, and the UAS is at a 300 ft altitude)
YOLOv8 Initial classifier (bounding box)	Classification thrashing over period of time	Person is detected, alternates between standing and laying several times in a short amount of time
YOLOv8 Initial classifier (bounding box)	Checking for object permanence vs. model hallucinations	A person is detected for very few frames in a certain time period, indicating a model hallucination
YOLOv8 Initial classifier and keypoint logic	Mismatch between logic and initial classifier	The keypoints say laying but the initial classifier said standing
YOLOv8 initial classifier	Measure of surprise and/or uncertainty based on training data	A bear is classified as human standing and receives a high surprise score
Key point classifier	Too few key points	We can't get a stance from these keypoints
Detection outputs	Calibrated Confidence Regression	Predicts calibrated confidence metric combining Intersection over Union (IoU, a metric used to measure how well a predicted bounding box overlaps with the ground truth bounding box) and classification correctness
Classification correctness	Correct Classification Predictor	Quantifies classification correctness using 2-layer BNN with Bernoulli output
Image inputs to models	Measuring characteristics of incoming data	An incoming image is washed out due to too much light

11 An Implementation of the RTA Framework

The RTA framework in this section represents a high-level description of the proof-of-concept prototype built during the execution phase of Aerial AId. The RTA framework was an asynchronous, multi-process platform for computer vision on an edge computing device (see Figure 34). The implementation was a system that combines computer vision components with an ensemble-based monitoring system to provide safety assurance for object detection and classification tasks. The platform is agnostic to the computer vision component being used; within the context of Aerial AId, it was planned to be used with the human stance classifier. YOLOv11 was the component used in testing the prototype. It was trained on the VisDrone dataset for detecting pedestrians and cars. The prototype was built in Python using a Continuous integration/Continuous development pipeline (CI/CD) in Gitlab so that versions of the model could be quickly developed and deployed as the hardware prototype became available, and updates could be made without introducing bugs in the code.

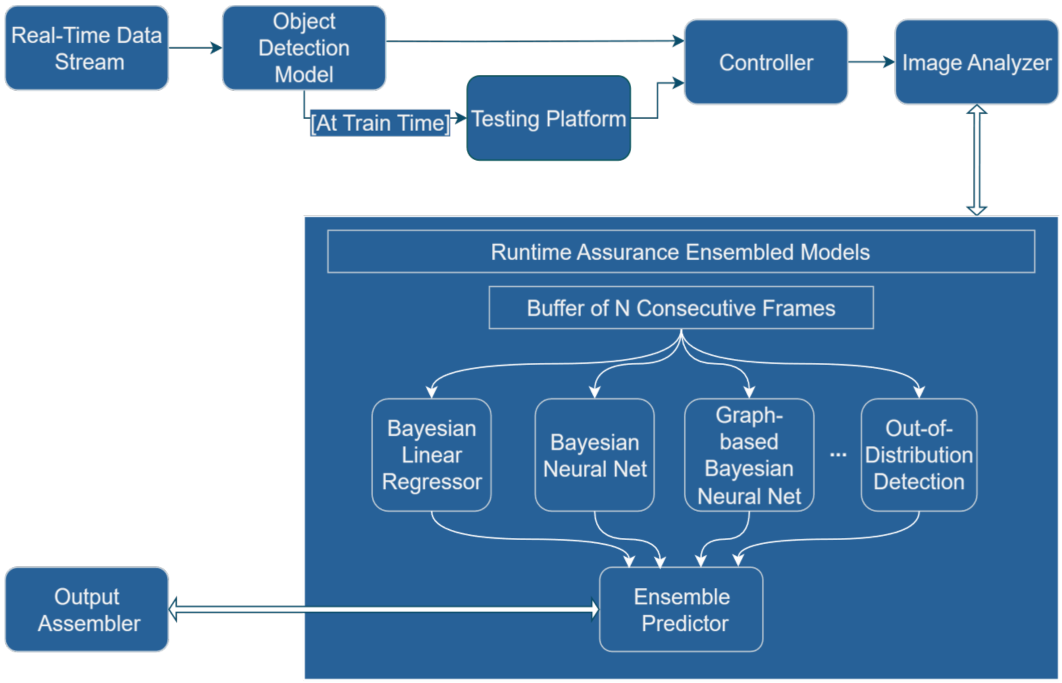

Figure 34: RTA framework architecture and data flows

11.1 System Architecture and Design

The RTA framework was built to be a modular, event-driven architecture. The system operated on the principle of separation of concerns, where each component maintains distinct responsibilities while participating in a coordinated data processing pipeline. This design philosophy enabled the platform to scale to the computational budget of the device it is running on. This design also offered the flexibility necessary for diverse application domains, where each component, including monitors and computer vision software, could easily be switched out.

The platform implemented a queue-based communication protocol. The communication system incorporated a fixed-size queue architecture, configured to hold two to four buffer entries to optimize the balance between system responsiveness and memory utilization. Each queue entry consisted of one buffer from the data processor, with each buffer containing 10 consecutive frames of data. This sliding window approach allowed for monitors that addressed temporal properties in the data and model performance, while maintaining efficient memory usage patterns.

Additionally, to support testing and validation of the runtime assurance platform, a testing infrastructure was developed that enabled experimentation and prototyping of monitors on data. The testing platform can input any YOLOv8 or YOLOv11 based model with a defined monitor to process a video segment. It can process the results of the computer vision component and monitors offline, or it can attempt to process results live and stream the results.

11.2 Data Processing

From the full images and individual detection bounding boxes, we extracted features that become possible inputs to our runtime assurance monitors. Table 6 describes the features extracted during data processing. In order to capture aspects of the video with respect to time, we explored various potential methods for time-series feature extraction. However, in the interest of computational efficiency for real-time performance, we utilize the simplest possible approach. By default, YOLO has the ability to track objects over frames, providing an object ID value. Given a sliding window buffer of frames for which we want to infer on the final frame, we pass forward the mean, standard deviation, and value in the final frame of each feature described above to the model to quantify uncertainties for each detection. Prior to model training, we normalized the numerical feature based data and did one-hot encoding for categorical features.

Table 6: Features extracted during data processing

SOURCE	DERIVED FEATURES	RAW FEATURES
Camera Frame	Contrast	N/A
	Entropy	
	Complexity	
	Spatial Complexity	
	Noise	
	Blur	
	Hue Variation	
	Shadow	
	Rain Streaks	
YOLO Predictions: Classification	N/A	Class Label
		Class Confidence
YOLO Predictions: Bounding Box	Aspect Ratio	Box Width
	Area Ratio	Box Height
	Edge Proximity	Normalized X, Y

11.3 Runtime Monitors

The RTA framework implemented is such that monitors can be developed and "plugged-in" easily. Several monitors were considered in this effort and mentioned in this section. Due to the time restriction, only a small number were implemented, leaving several monitor implementations for future work.

One class of monitors utilized Bayesian Neural Networks (BNNs), which are capable of uncertainty quantification. While traditional machine learning models are trained to find optimal deterministic parameters for a function that maps the inputs to the outputs of the training data set, Bayesian models treat each parameter and the output as a distribution [54, 55]. Thus, Bayesian models output many samples for each input, which produces a point mean prediction but also a standard deviation that can be used for uncertainty quantification.

Several models of varying complexity were tested with different objectives. This included the three listed below.

- **Calibrated Confidence Regression**. This monitor is trained to predict a custom "Calibrated Confidence" (CC) metric for each detection, defined using the following equation:

  $$CC = IoU * 1_{\text{prediction class==ground truth class}},$$

  where IoU is the "Intersection over Union" ratio of the number of pixels in the intersection of the predicted and true bounding boxes with the number of pixels in the union of the predicted and true bounding boxes. The model is a Beta regression with a 2-layer BNN for the alpha and beta values that parameterize the output Beta distribution, fit with Stochastic Variational Inference (a machine learning technique for approximating probability distributions).

- **IoU Regression**. This monitor focuses on quantifying localization uncertainty of the object detection. It consists of a 2-layer BNN with Gaussian output distribution, fit using Markov chain Monte Carlo (MCMC) sampling.

- **Correct Classification Predictor**. This monitor focuses on quantifying classification correctness. It consists of a 2-layer BNN with Bernoulli output distribution, fit using MCMC sampling.

The ensemble of monitors feeds into the output module, which communicates information to the human operator. We communicate information through four main means:

- **Aggregate Uncertainty**- For each detection, we display the uncertainty of the detection alongside the default YOLO confidence as a value normalized between 0 and 1. This value is taken as a weighted average of the outputs of the calibrated confidence regression, IoU regression, and correct classification predictor, aggregating information on localization and classification uncertainty into a single value.

- **Localization Uncertainty**- For each detection, we display the localization uncertainty as the transparency of the output bounding box.

- **Classification Uncertainty**- For each detection, we display the classification uncertainty as the color of the output bounding box on a gradient from red to green, with red being highly uncertain and green being low uncertainty.

11.4 Example of Outputs

A working version of the RTA framework is available in Python. There is still significant work to be done developing monitors, but the system allows different monitors to be switched in and out, as well as different computer vision components. Figures 35–36 show two examples of what the output of the RTA framework looks like. Figure 37 shows Aerial AId summer interns prototyping the RTA framework on a live stream of data.

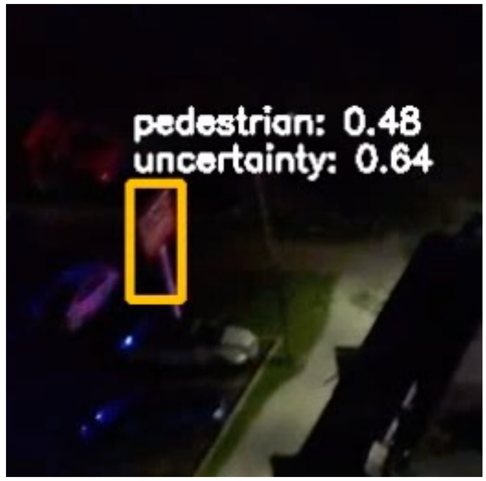

Figure 35: RTA framework classifying restaurant sign as pedestrian[7]

Figure 36: RTA framework classifying cars[7]

[7]Original image source: BCFPD/MO-TF1 DSAR team, https://dsar.motf1.com/

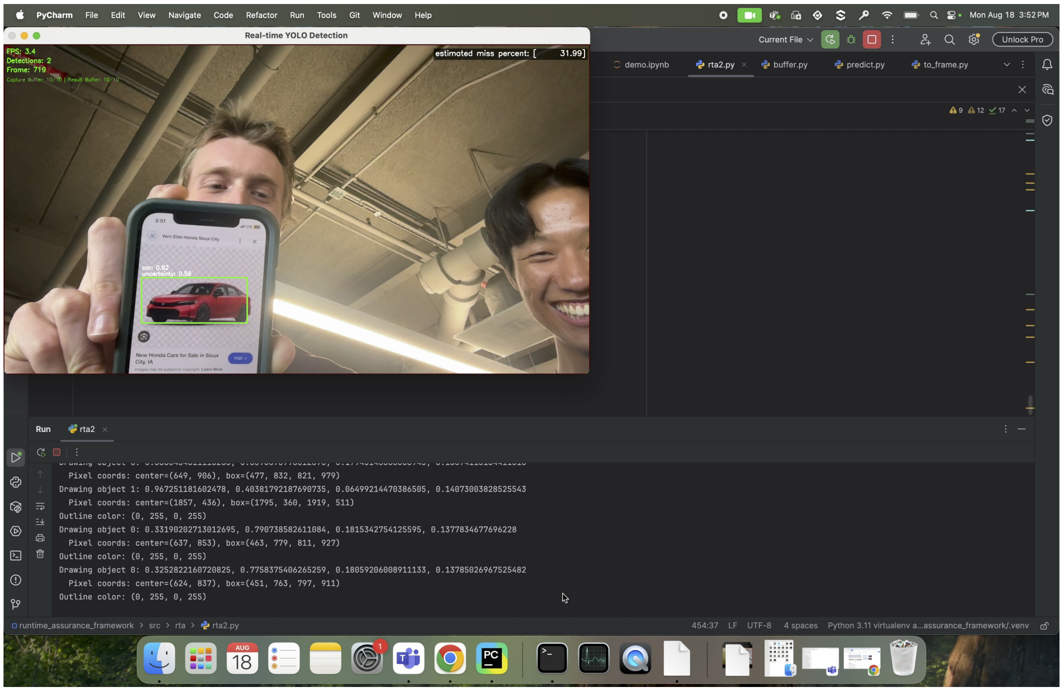

Figure 37: Summer 2025 interns prototype the RTA framework with live video feed

12 Applying Formal Methods Towards Perception Engines

12.1 Formal Specification of Convolutions

Formal verification is a strong form of assurance that involves the application of mathematics and logic to software and hardware systems. Interactive theorem provers are one tool used to perform such verification, and NASA has a long history of using SRI International's Prototype Verification System (PVS) [56] for this kind of work [57]. PVS allows users to describe a system and its safety properties within a specification language and provides a proof-theoretic environment in which to prove that the system satisfies those properties. These methods have been used successfully by the NASA Langley Formal Methods Research Group to verify safety-critical systems in aeronautics [58, 59], and as a platform to develop specialized tools for novel problems [60].

The problem of applying formal methods to machine learning is recognized as a difficult one [61]. The statistical nature of the tools which are commonly applied to the technology creates a foundation upon which to build assurance claims to a high enough confidence level for use in safety-critical systems. Attempts to work around this often rely on over-approximation or on modeling neural network structure at a very low level [62]. The Aerial AId project's choice of using convolutional neural networks for object detection presented a different opportunity.

The same structure of a convolutional neural network which allows for its efficient execution also allows for a mathematical generalization of its function. Furthermore, Commercial Off-The-Shelf (COTS) solutions such as YOLO exhibit higher level repetition by chaining similar components together into a "backbone" [63]. By focusing on these components and their intended function, the team was able to formally model a discrete convolutional operation in PVS. Much of the computational content in a convolutional neural network is performed by these convolutional operations, and the key "learnable" content in each convolution is the kernel. These kernels provide a much more tractable target for verification. Where a production neural network may involve millions of neurons, a typical kernel might be a 3×3 matrix [63].

Formal verification is often performed either prior to development with design-time algorithm verification, or after development with code verification. The structure of Aerial AId more closely resembled an iterative, agile development process. Integrating formal methods into such processes is an active area of interest. The team mirrored their formal model in a programming language (Typed Racket [64]) which allowed close syntactic correspondence with the PVS specification language. This allowed the team to quickly cycle between comparison, testing, and proving in a manner which promoted rapid refinement of the formal models.

12.2 Modeling RTA Systems as Hybrid Programs in Plaidypvs

The team developed a formal verification framework for runtime assurance systems using Plaidypvs (**P**roperly **A**ssured **I**mplementation of **D**ifferential Dynamic Logic for H**y**brid **P**rogram **V**erification and **S**pecification). Plaidypvs is a formal methods tool for the reasoning of hybrid systems [65]. In this framework, it is possible to define an RTA system and prove general requirements about what it takes for the RTA system to satisfy a certain safety property. This work established foundational mathematical proofs for the safety properties of simplex RTA architectures, providing a basic approach that could be extended to more complex runtime assurance frameworks like those developed for computer vision systems in Aerial AId.

The formal verification approach modeled the simplex RTA framework as a hybrid system using differential dynamic logic, where both the advanced controller (AC) and reversionary controller (RC) are represented as hybrid programs α and β, respectively. The complete RTA system can be expressed as the hybrid program:

$$((?M; m_{\tau,M}(\alpha)) \cup (?\neg M; \beta))^*$$

where M represents the monitor condition that triggers switching between controllers, and $m_{\tau,M}$ ensures that the monitor is checked at least every τ time units. This formalization captures the essential behavior of runtime assurance: the system operates under the advanced controller α as long as the safety condition M holds, but switches to the trusted reversionary controller β when $\neg M$ is violated.

The framework includes a proven rule for establishing safety properties of the overall RTA system. The fundamental RTA rule demonstrates how the safety of the complete system can be decomposed into safety properties of its individual components:

$$\frac{\Gamma \vdash S \wedge (M \vee G) \quad S \vdash [m_{\tau,M}(\alpha)](S \wedge (G \vee M)) \quad G \vdash [\beta^*]S}{\Gamma \vdash [((?M; m_{\tau,M}(\alpha)) \cup (?\neg M; \beta))^*]S} \ (\mathbf{RTA})$$

This rule decomposes the complex safety proof of the entire RTA system into three subgoals, each addressing a specific aspect of system behavior. The first subgoal, $\Gamma \vdash S \wedge (M \vee G)$, establishes that the system begins in a safe state where either the monitor condition holds, or a user-defined switching condition, G, is satisfied. The second subgoal, $S \vdash [m_{\tau,M}(\alpha)](S \wedge (G \vee M))$, proves that the advanced controller maintains safety and appropriate switching conditions throughout its monitored execution. The third subgoal, $G \vdash [\beta^*]S$, ensures that the reversionary controller can maintain safety indefinitely once activated. By proving these three component-level properties, the rule guarantees that the overall RTA system will maintain the safety property, S, throughout its operation.

While this formal verification work focused on relatively simple UAS flight control examples, the general framework it established is directly applicable to the more sophisticated runtime assurance systems developed for Aerial AId. The computer vision monitoring framework described in previous sections represents a natural extension of the simplex architecture. The formal verification principles demonstrate that such monitoring systems can provide mathematically rigorous safety guarantees, even when the primary perception system relies on complex AI/ML algorithms that cannot be directly verified.

Part V
Hardware Specification and Requirements

To achieve Aerial AId's vision of a fully autonomous DFR system, it is necessary to implement, test, and evaluate specific YOLO algorithm variants under realistic operational conditions, while simultaneously collecting rich datasets of images and sensor data. Off-the-shelf UASs typically do not expose low-level access to sensor data or allow modular customization of processing architecture, limiting experimentation with novel models, data pipelines, or optimization strategies. Instead, the Aerial AId team decided to engineer a custom UAS payload, which provides complete control over all hardware components, from selecting the precise sensors best suited for the specific application scenario to integrating a dedicated edge computation unit capable of executing the expected YOLO models onboard in real time. The payload design was guided by the following set of internal requirements and operational constraints aimed at supporting onboard YOLO-based human stance detection while maintaining broad UAS compatibility:

- The payload needed to capture and store aerial imagery of sufficient quality for real-time YOLO pose classification and to run the AI algorithms directly on dedicated edge computing hardware.

- To maximize flexibility and reduce integration complexity, the payload was required to be fully self-sustained, relying on its own power supply, computation, and communications rather than those of the host UAS.
- The payload was to be designed to be UAS-agnostic, with an initial target mass under 10 lb to ensure compatibility with a wide range of airframes. This target mass was later increased to account for heavier payload configurations such as those using a Jetson AGX Orin with a stabilized gimbal.
- Nominal operation was defined at an altitude of 30 m (approximately 100 ft), while maintaining flexibility for higher or lower altitudes as dictated by mission conditions.
- To support longer data collection, the system was required to sustain at least 2 hours of continuous operation on a dedicated battery set.

Given these requirements, the Aerial AId team started the design process by evaluating a range of potential hardware configurations that could satisfy both the sensing and computational demands while meeting weight and operational constraints. Early discussions led to the definition of two baseline approaches: a minimal configuration, representing the minimum viable hardware configuration required to meet the project's objectives (in a non-optimal, but functional way), and a desired configuration, which incorporated additional sensors, payload stabilization, and remote control and monitoring capabilities.

The minimal configuration consisted of a visible-light camera, an onboard edge computing device, a dedicated power source, and a UAS mounting bracket. This bare minimum setup provided the capability to collect and store visible-spectrum aerial imagery while performing real time YOLO classification on edge hardware. However, several tradeoffs accompany this approach. Payload software control (e.g., starting or stopping data logging) is only possible while the UAS is on the ground, before take-off or after landing. Similarly, output data can only be viewed offline after the flight, with no real-time feedback available to the payload operators. Additionally, the basic mounting bracket does not offer any image stabilization and the limited sensor suite restricted imagery to the visible spectrum with no geolocation capability.

The desired configuration expanded the payload's capabilities by also incorporating a stabilized gimbal for precise payload pointing and image stabilization, an infrared camera to extend sensing into the thermal spectrum, and a GNSS/INS (Global Navigation Satellite System/Inertial Navigation System) unit to enable software image stabilization and provide accurate geolocation of collected imagery. In addition, an onboard cellular modem leveraged the existing commercial cellular network to provide bidirectional data relay, enabling operators to monitor payload status, adjust software settings, and receive classification results in real time. A schematic overview of this desired configuration is shown in Figure 38.

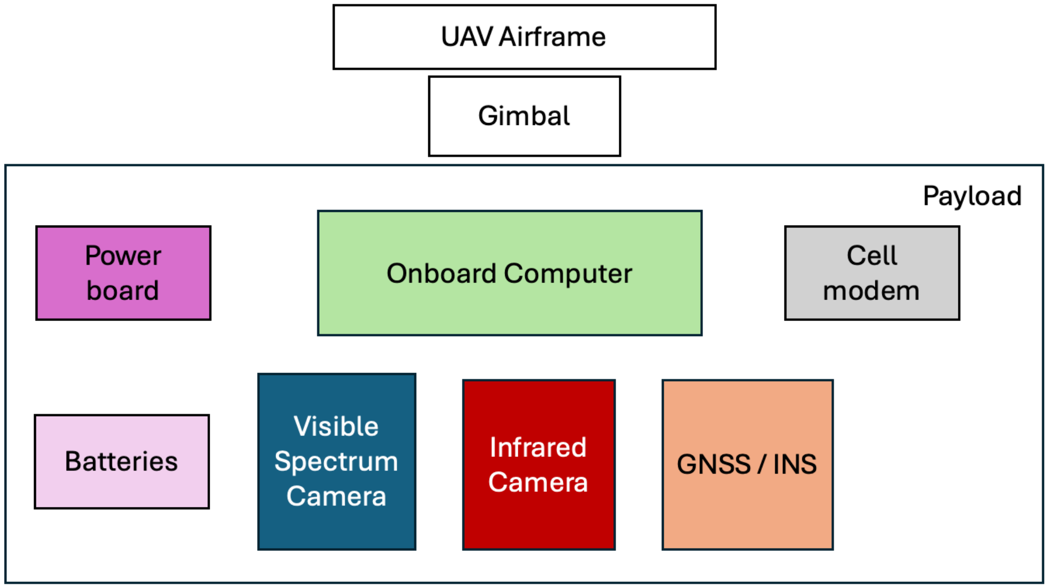

Figure 38: Schematic overview of the desired hardware configuration for the Aerial AId payload

Given the desired hardware configuration described above, the team started to survey existing commercial devices that could satisfy the identified requirements. This process focused on evaluating available off-the-shelf solutions that offered the necessary functionality while maintaining interface compatibility with the other components and fitting within the payload's SWaP (Size, Weight, and Power) budget.

12.3 Onboard Computer Selection

The selection of the onboard edge computing module was primarily driven by the requirements of the project's AI-development team, who were already developing and testing YOLO models on the NVIDIA Jetson AGX Orin platform. Using this hardware in the payload would ensure seamless integration of their existing workflows, with the main drawback being its high SWaP, with a size of 110 × 110 × 72 mm, a mass of 880 g, and a power consumption of up to 60 W. Less capable Jetson devices, such as the Jetson Orin Nano Super, were identified as potential fallback options that could reduce the SWaP demands (100 × 79 × 34 mm, 320 g, 15 W). However, the Aerial AId's AI team noted that their models (YOLOv8s) could not run on these platforms without significant optimization. As a result, the Jetson AGX Orin was adopted as the preferred solution for the desired hardware configuration. This edge computer offers the necessary AI performance while providing large data storage on an external M.2 Solid State Drive and sufficient Input/Output (I/O) interfaces to connect the desired sensors and devices.

12.4 Visible Spectrum Camera Selection

In parallel with selecting the computing module, the team initiated a survey of commercial visible-spectrum cameras that could meet the payload's imaging requirements. Compatibility with NVIDIA Jetson boards was a critical constraint to ensure seamless integration with the chosen onboard computer. Aerial AId's AI team recommended selecting a camera with at least a full HD resolution and 30 FPS, aligning the payload's imaging configuration closely with common specifications of integrated commercial UAS cameras. This approach would facilitate potential future integration of the Aerial AId algorithms directly with commercial UAS hardware. Additionally, team members with prior experience in collection of airborne imagery suggested selecting a camera with global shutter to avoid rolling shutter artifacts caused by motion and airframe vibrations. With these requirements in mind, the team evaluated different types of camera interfaces, focusing on the more common USB, GigE, and MIPI interfaces. USB cameras are widely available and supported but are often limited in bandwidth depending on their USB interface (5 Gbps with USB 3.0, 10 Gbps only with the less adopted USB 3.2 Gen 2). GigE (Gigabit Ethernet) support long cable runs and multiple cameras on one network (properties that are not particularly useful for the Aerial AId payload) but have a lower bandwidth of 1 Gbps (10 Gbps for the less common 10GigE cameras) and can experience higher latency. MIPI CSI-2 (Mobile Industry Processor Interface, Camera Serial Interface 2) is a high-speed, low-power standard for transmitting image/video data in embedded devices which offers low latency data transfer of up to 10 Gbps. Despite only supporting short cable lengths (usually less than 30 cm) and requiring dedicated hardware interfaces, this interface was chosen as the preferred one for the Aerial AId payload, with USB being a potential fallback option.

The NVIDIA Jetson developer website provides a catalog of camera vendors and models verified for Jetson compatibility. By narrowing the search to devices supporting the MIPI CSI-2 interface, equipped with global shutter camera sensors, offering driver compatibility across multiple Jetson boards (AGX Orin/Xavier, NX, TX2, Orin Nano), and supplied by vendors compliant with Section 889 of the FY19 National Defense Authorization Act (NDAA), Allied Vision emerged as a suitable choice. Among several Allied Vision products that met the desired specifications, the team selected the Alvium 1800 C-811c model as the optimal choice. This MIPI CSI-2 camera body is equipped with a Sony IMX646 CMOS, Type 2/3, global shutter sensor that can deliver a resolution of 2848×2848 pixels (8.1MP) at 59 FPS. Both the resolution and frame rate can be changed via software by selecting a different Region of Interest (ROI). The camera body is $26 \times 29 \times 29$ mm in size and weighs 40 g, with a typical power consumption of 3.1 W, taken directly from the MIPI CSI-2 interface. Allied Vision offers an adapter board to connect the Alvium 1800 C-811c camera to a Jetson AGX Orin. Special care is needed for the selection of the FPC (Flexible Printed Circuit) cable between the camera body and the adapter board, with the standard FPC cable and retaining plate only recommended for evaluation purposes and the Flex FPC cable suggested for drone and airborne applications.

The lens selection process balanced wide field of view for increased situational awareness with sufficient zoom for detailed visibility. Applying imaging fundamentals to the expected operational scenario, as shown in Figure 39, leads to the following formulas for calculating the viewable area and pixel density:

1. Viewable area: $h \approx z \cdot r \cdot s/(1000 \cdot f \cdot sin(\theta_{tilt}))$
2. Pixel density: $px_d \approx 1000 \cdot f \cdot sin(\theta_{tilt})/(z \cdot s)$

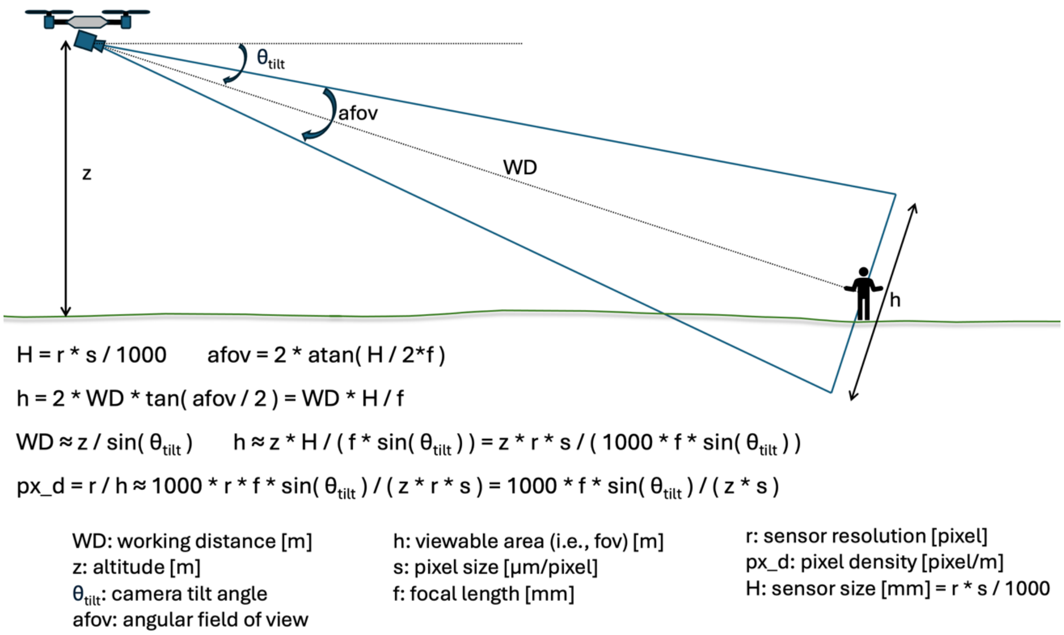

Figure 39: Imaging fundamentals applied to expected operational scenario to calculate expected viewable area and pixel density

In these formulas, the sensor resolution (r), and pixel size (s), are properties of the camera body; the altitude (z), and tilt angle (θ_{tilt}), depend on the operational scenario; the focal length (f) is a property of the lens, which can be constant (fixed focal length lenses) or variable (zoom lenses).

Figure 40 shows a plot of viewable area (left axis) and pixel density (right axis) as a function of the lens focal length in millimeters, calculated for the Allied Vision Alvium 1800 C-811c sensor specifications. Based on more recent interpretations of Johnson's criteria for evaluating the performance of imaging systems, reliable AI object identification requires at least 24 pixels across the target. Approximating that to a pixel density of 24 pix/m allows the identification of a viable range of focal lengths for the proposed scenario that balance the viewable area with enough pixel resolution.

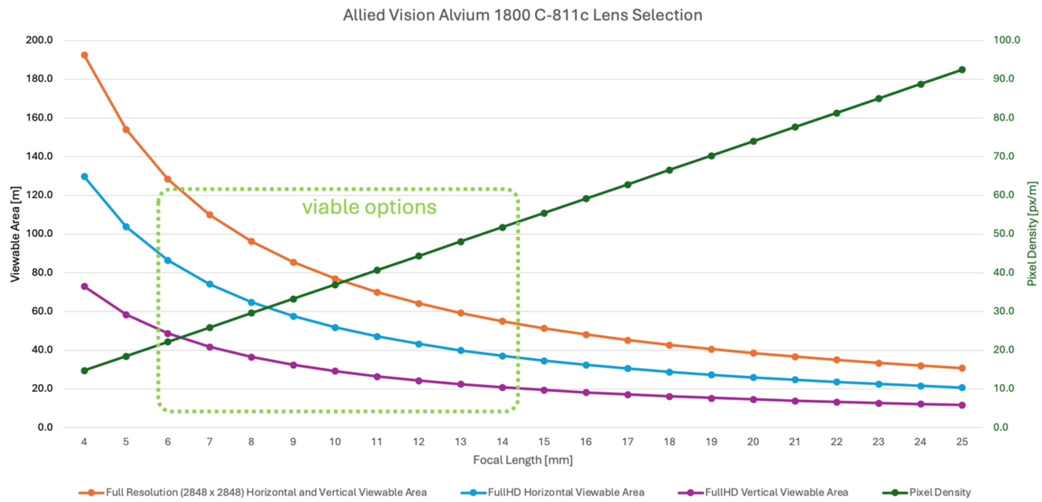

Figure 40: (Left axis) viewable area in meters at different camera resolutions. (Right axis) pixel density (pixels/m) as a function of lens focal length

While a final lens selection has not been made, a potential candidate that meets the conditions described above is the Edmund Optics TECHSPEC HRr Series 12mm f/4 lens. This fixed focal length ruggedized lens, in which all the individual lens elements are glued in place to reduce pixel shift, is recommended for machine vision and object tracking applications. Combined with the Allied Vision Alvium 1800 C-811c camera body at full resolution, this lens provides a viewable area of 64×64 m (210×210 ft) with a pixel density of 44 pix/m, sufficient to demonstrate feasibility of the Aerial AId concept.

12.5 Infrared Camera Selection

A similar analysis was carried out for the thermal imaging component of the payload. The primary purpose of this sensor within the Aerial AId payload is the detection of human body heat signatures, with the added potential to provide a rough estimate of body temperature, which can help improve the quality of the scene assessment and medical triage. Among the infrared imaging bands: Near Infrared (NIR), Mid-Wave Infrared (MWIR), Long-Wave Infrared or (LWIR), the team identified the LWIR spectrum as the appropriate choice. LWIR cameras are directly sensitive to emitted body heat, do not require external illumination, work both in daylight and at night, and maintain robust performance in degraded visual environments such as smoke or fog. Although LWIR sensors generally exhibit lower sensitivity compared to MWIR systems, the widespread availability of uncooled LWIR detectors provides significant advantages in terms of size, weight, power, and cost, avoiding the complexity and expense associated with cryogenically cooled MWIR alternatives.

Within this context, the team surveyed several commercial vendors, evaluating any documented compatibility with the NVIDIA Jetson boards and their Section 889 compliance, and identified the BOSON+ series from Teledyne FLIR as a suitable option for integration. These LWIR cameras offer a 640×512 pixel resolution with a sensitivity of less than 30 mK for their professional-grade model and a frame rate of 60 Hz. They support both USB3 and MIPI interfaces, with documented examples available online demonstrating integration with NVIDIA Jetson board. The BOSON+ camera body is very compact and lightweight, with a size of $21 \times 21 \times 11$ mm (without lens), a mass of 7.5 g (also without lens), and a power consumption of approximately 1 W. The camera is available with a factory-installed fixed lens, with options offering horizontal angular fields of view ranging from 6 degrees to 95 degrees. By applying the same formulas used to down select visible-spectrum cameras, the team calculated the viewable area and pixel density for each lens option, as shown in Table 7. The table shows the viewable area and pixel density for each Boson+ 640 lens configuration, taken from an operational altitude of 100 ft and tilt angle of -18 degrees, with the viable options highlighted in green that offer enough pixel density.

Table 7: Viewable area and pixel density for each Boson+ 640 lens configuration

Horizontal AFOV [deg]	95°	50°	32°	24°	18°	12°	8°	6°
Horizontal Viewable Area [m]	215.4	92.0	56.6	42.0	31.3	20.7	13.8	10.3
Vertical Viewable Area [m]	172.3	73.6	45.3	33.6	25.0	16.6	11.0	8.3
Pixel Density [px/m]	3.0	7.0	11.3	15.3	20.5	30.8	46.4	61.9

Comparing the table with the curves in Figure 40 shows that the lower resolution inherent to IR cameras strongly affects the tradeoff between viewable area and pixel density, reducing the achievable viewable area to roughly one-fifth that of a visible-spectrum camera at similar pixel density. Therefore, the BOSON+ 640 with an 18 deg lens is recommended for Aerial AId, offering the largest viewable area at the minimal pixel density necessary for reliable object identification.

12.6 Gimbal Selection

A UAS-mounted gimbal plays a critical role in enabling reliable sensing and imaging. At the most basic level, the gimbal stabilizes onboard sensors, maintaining their desired pose even as the UAS experiences vibration or motion. This stabilization reduces image jitter and motion blur, thereby improving the quality of collected data. Gimbals can also be used for active pointing, directing sensors toward a specific target of interest and keeping that target centered in the field of view despite motion of either the UAS or the target itself. Such functionality could be achieved either manually, through operator joystick control, or automatically via software commands that close the loop between detection, tracking, and gimbal actuation.

In evaluating gimbal options, two main design approaches were considered. A smaller gimbal that stabilizes only the cameras offers advantages in SWaP but introduces complexities in integrating within the Aerial AId payload and routing cables from the fixed computer to the moving sensors. In contrast, a larger gimbal capable of stabilizing the entire Aerial AId payload is heavier but more flexible, as it is less complicated to integrate and allows easier replacement with alternative models and the option to fly the payload without a gimbal at all. Given these tradeoffs, the larger gimbal approach was deemed preferable, despite exceeding the initial 10 lb payload allocation. A viable candidate is the FREEFLY MōVI Pro Aerial Bundle, which the team was able to assess due to its use on other NASA projects. This system weighs 2.7 kg, can stabilize payloads up to 196 × 200 × 175 mm in size and 6.8 kg in mass, and is powered by a dedicated set of batteries. Furthermore, its serial interface can be accessed by the NVIDIA Jetson for software-driven control of gimbal pointing angles, enabling both manual and automated targeting capabilities.

12.7 Other Components Selection

In addition to the gimbal, onboard computer, and imaging sensors, the remaining payload components were primarily chosen from hardware that the team already had access to or had previous experience integrating into UAS payloads. For the GNSS/INS unit, the recommended solution is the VectorNav VN-200. This compact, high-performance sensor combines a GPS receiver with a tactical-grade MEMS IMU in a single unit, providing position, velocity, and attitude estimates in real time. Its small form factor (25 × 25 × 3 mm) and low power consumption (0.5 W) make it well suited for UAS payload integration.

For communications, a Verizon cellular modem is available and can be used to transmit telemetry and data through the mobile network. While this provides a convenient means of connectivity during field tests in areas with strong coverage, it is not considered a long-term robust solution for deployment in emergency scenarios, where the cellular network may be unstable or entirely unavailable.

Power for the payload is supplied by a Lumenier 12000 mAh 4s 20C LiPo battery, which is expected to provide roughly two hours of continuous runtime under nominal load conditions. Power conversion to the necessary regulated voltage is handled by an off-the-shelf 12 V, 72 W DC-DC converter from SparkFun Electronics, ensuring stable power delivery to all payload components.

As a candidate housing design, the team considered a structure composed of custom T-slot aluminum extrusions, carbon fiber plates, and a molded plastic outer shell. This configuration is intended to ensure a highly rigid and stable connection between all payload components, minimizing any relative shift that could compromise prior sensor calibrations. In addition, it provides a secure interface for mounting to the external gimbal and offers basic protection against environmental exposure.

In summary, the desired payload hardware configuration can be assembled from the following set of commercially available components:

- NVIDIA Jetson AGX Orin 64 GB
- Allied Vision 1800 C-811c camera
- Allied Vision adapter board for AGX Orin + connection cables
- Edmund Optics TECHSPEC HRr Series 12mm f/4 lens
- Teledyne FLIR BOSON+ 640, Professional, Shuttered, 18 deg lens + USB kit
- VectorNav VN-200 Rugged Development Kit

- FREEFLY MōVI Pro Aerial Bundle
- Verizon cellular modem
- Lumenier 12000 mAh 4s 20C LiPo battery
- SparkFun Electronics 12 V, 72 W DC-DC converter

Based on estimates at the time of this report, the complete desired hardware configuration, including gimbal, onboard computer, sensors, power system, communication hardware, and housing, would amount to an approximate total cost of $20,000 USD. Alternatively, a minimal configuration consisting only of the enclosure, onboard computer, power system, and visible camera, could be assembled for approximately $5,000 USD.

13 Discussion and Future Work

Aerial AId was planned to be a two-year CAS execution activity but went into an orderly closeout after nine months due to the ending of the CAS project within ARMD. Partially because of this, there are several tracks of work that could be picked up and continued. The significant achievements of Aerial AId were:

- Two sets of data collection flights for capturing human stance data from an aerial perspective. This includes a machine learning dataset, and even more raw data
- A tool for generation of synthetic emergency environments in the Unity engine
- A nominal human stance classifier that utilizes keypoint extraction and a logic to deduce the stance of a human on the scene
- An implementation of a runtime assurance framework for an ML/AI computer vision component
- Hardware requirements for an edge computing device to operate the runtime assurance framework complete with computer vision components
- The beginnings of a convolution library in a theorem prover, for the formal specification and reasoning of computer vision models
- A formal verification effort for runtime assurance systems that allows black box assurance of AI/ML components

In addition to these technological achievements, the team has built relationships with medical UAS stakeholders including DRONERESPONDERS, Brookhaven PD, The Matador Consortium, Missouri Task Force 1, FEMA, The Langley Joint Task Force - Civil Support, BRINC Drones, and ParaWave. These relationships were built over time and hold immense value beyond the Aerial AId activity.

There are still several avenues for future work, the most apparent being the scheduled second half of the Aerial AId execution. This would include:

- Building the edge-computer hardware prototype for the runtime assurance framework, including benchmarking through a series of lab tests and eventual integration onto a UAS as a research payload
- The iterative development of the runtime assurance framework, including integration of the human stance classifier prototype and more monitor integration, including geometric feasibility monitors and surprise-adequacy-based monitors
- Further processing of raw data from the test flight into a larger machine learning dataset focused on human stance, with a nominal human stance classifier trained on the data
- A flight demonstration day with the completed edge assurance framework prototype flying over a mock emergency scenario
- Creating and benchmarking a more robust classification system for human stance

14 Conclusions

Aerial AId represented an activity that was the culmination of design thinking for developing research efforts for the benefit of humanity. The approach brought in hundreds of stakeholders and emerged from a process that relied on an uncounted number of NASA civil servants and contractors. The result of this process was a window into what the future of agencies like NASA are capable of: A multi-disciplinary, multi-center team spanning early career and seasoned subject matter experts, collaborating to execute a vision they were all responsible for and which directly tied into the needs of humanity. In this case, the vision was directly traceable to the needs of first responders that work tirelessly to provide safe and timely care to people in need.

The CAS project built the infrastructure to make the discovery process happen. CAS built the room and brought in the facilitators, giving NASA civil servants a space to create without having to rely on traditional means of project creation which can be rigid and unnecessarily dependent on authority and precedent. Through this approach, Aerial AId developed not only technical innovations, such as a runtime assurance framework and synthetic dataset, but also represented a methodology for stakeholder-driven research that directly addressed real-world operational needs. With the closing out of CAS, this infrastructure will go away. Inside and outside of the agency, the workforce that contributed to CAS, along with the collaborative approaches and stakeholder relationships they cultivated, offers the largest transition opportunity to future efforts. As long as there are people committed to solving the hard problems for the benefit of all, they will find ways to create progress.

Appendices

A Executive Summary

The Aerial AId activity was a NASA-led research initiative under the Convergent Aeronautics Solutions (CAS) program, conducted from October 2024 through early 2025. The activity was developed through the CAS discovery process, an agile development approach that tightly integrated NASA researchers and relevant stakeholders in the emergent UAS medical industry through co-design sessions and workshops. The project consisted of a team of civil servants and contractors from NASA's Langley, Glenn, Ames, and Armstrong Research Centers. The project was designed to enable fully autonomous Drone-as-a-First-Responder (DFR) capabilities by developing trustworthy computer vision systems for situational awareness of emergency medical scenes. Originally planned as a three-year effort through FY26, Aerial AId was terminated approximately one year early due to the conclusion of the broader CAS program.

A.1 Mission and Objectives

Aerial AId pursued three core technical objectives:

1. Creating specialized datasets for computer vision system development
2. Developing nominal computer vision systems for rapid emergency scene assessment, and
3. Establishing runtime monitoring systems to assure AI/ML component performance

The project directly addressed critical barriers in emergency medical response by developing technology to improve situational awareness, reduce response times, and decrease operational strain on human first responders.

A.1.1 Key Accomplishments

Despite the shortened timeline, Aerial AId achieved the following significant technical milestones:

1. Successfully conducted data collection flight tests at the CERTAIN I test center and Kennedy Space Center in collaboration with the AIRVUE Activity within Transformational Tools and Technologies, collecting approximately 500 GB of aerial video data focused on human stance detection in emergency scenarios.
2. Developed the Synthetic Aerial View (SAiV) tool, a novel synthetic data generator built on Unity 3D and the PeopleSansPeople High Dynamic Range Image framework, specifically designed for privacy-preserving aerial human detection and pose estimation tasks.

3. Created a nominal human stance classifier that utilizes keypoint extraction and a logic to deduce the stance of a human on the scene, based on the You Only Look Once (YOLO) object detection and classifier.

4. Documented hardware requirements for an edge computing device to operate the runtime assurance (RTA) framework complete with computer vision components.

5. Developed the beginnings of a convolution library in a theorem prover, for the formal specification and reasoning of computer vision models.

6. Developed a formal verification effort for runtime assurance systems, that allows reasoning of AI/ML components.

7. Developed an RTA framework featuring an asynchronous, multi-process runtime assurance platform with ensemble-based monitoring capabilities for safety-critical computer vision applications, implemented in Python.

8. Completed a market research report analyzing the state of uncrewed aerial systems for emergency medical first response, providing industry insights and technological assessments that informed the project's technical approach and stakeholder engagement strategy.

A.2 Teaming and Stakeholder Engagement

Aerial AId successfully established partnerships across the emergency medical response ecosystem, including medical sUAS industry startups, emergency medical service providers, search and rescue operators, and the DRONERESPONDERS organization. The project leveraged existing collaborative frameworks, including access to tornado video data from Missouri Task Force 1 through an established Space Act Agreement, and engaged with the Langley Joint Task Force - Civil Support to strengthen emergency response capabilities. The multicenter NASA team, comprising scientists from Langley, Glenn, Armstrong, and Ames Research Centers, demonstrated effective collaboration across AI/ML data science, ML assurance, and flight test disciplines.

A.3 Future Work

The project's deliverables provide a strong foundation for continued development through alternative funding mechanisms. Key technologies ready for transition include the runtime assurance framework for computer vision applications, the SAiV synthetic data generation tool, the processed aerial ML dataset focused on human stance detection, and preliminary vision models for emergency scene assessment. These technologies have immediate applications beyond emergency medical response, including search and rescue operations, disaster response, and broader autonomous aerial vehicle missions.

The early termination of Aerial AId, while premature, does not diminish the technical progress achieved or the critical need for this technology in emergency response applications. The team recommends pursuing continuation of this work through alternative NASA programs, partnerships with other government agencies, or industry collaborations. Priority should be given to completing the dataset preparation for public release, further developing the runtime assurance framework, and establishing technology transfer agreements with interested stakeholders. The foundation established by Aerial AId provides a pathway for achieving the original vision of fully autonomous emergency medical response capabilities.

The Aerial AId activity worked to demonstrate the feasibility of trustworthy computer vision for emergency medical aerial systems and established the technical and stakeholder foundations necessary for continued development of this critical technology.

B List of Acronyms

AC	Advanced Controller
ACERO	Advanced Capabilities for Emergency Response Operations
AFRC	NASA Armstrong Flight Research Center
AIRVUE	Airborne Instrumentation for Real-world Video of Urban Environments
AOSP	Airspace Operations and Safety Program
APNT	Alternative Position Navigation and Timing
ARC	NASA Ames Research Center
ARMD	Aeronautics Research Mission Directorate
ASTM	(formerly known as) American Society for Testing and Materials
ATD	Anthropomorphic Testing Device
ATM-X	Air Traffic Management – eXploration
BCFPD	Boone County Fire Protection District
BNN	Bayesian Neural Networks
CAS	Convergent Aeronautics Solutions
CC	Calibrated Confidence
CERTAIN	City Environment Range Testing for Autonomous Integrated Navigation
CI/CD	Continuous Integration/Continuous Development
CNN	Convolutional Neural Networks
COCO	Common Objects in Context
COTS	Commercial-off-the-Shelf
DARPA	Defense Advanced Research Projects Agency
DFR	Drone-as-a-First-Responder
DL	Deep Learning
DSAR	Disaster Situational Assessment and Reconnaissance
EMS	Emergency Medical Services
EO	Electro-Optical

FAA	Federal Aviation Administration
FEMA	Federal Emergency Management Agency
FPC	Flexible Printed Circuit
FY	Fiscal Year
GNSS	Global Navigation Satellite System
GRC	NASA Glenn Research Center
I/O	Input/Output
INS	Inertial Navigation System
IoU	Intersection over Union
IR	Infrared
LaRC	NASA Langley Research Center
LiDAR	Light Detection and Ranging
LWIR	Long-Wave Infrared
MCMC	Markov chain Monte Carlo
ML	Machine Learning
MO-TF1	Missouri Task Force 1
MWIR	Mid-Wave Infrared
NASA	National Aeronautics and Space Administration
NDAA	National Defense Authorization Act
NIH	National Institute of Health
NIR	Near Infrared
NIST	National Institute of Standards and Technology
NOMAD	Natural, Occluded, Multi-scale Aerial Dataset
NSF	National Science Foundation
OCIO	Office of the Chief Information Officer
OCR	Opportunity Concept Report
OGA	Other Government Agency
OOTB	Out-Of-The-Box

P3	Persevere, Pivot, or Punt
Plaidypvs	Properly Assured Implementation of Differential dynamic logic for hYbrid Program Verification and Specification
PSP	PeopleSansPeople
PSP-HDRI+	PeopleSansPeople High Dynamic Range Image
PVS	Prototype Verification System
RC	Reversionary Controller
RGB	Red, Green, Blue
ROI	Region of Interest
ROS	Robotic Operating System
RTA	Runtime Assurance
SAiV	Synthetic Aerial View
SWaP	Size, Weight, and Power
SWS	System-wide Safety
TACP	Transformative Aeronautics Concepts Program
TTT	Transformational Tools and Technologies
UAM	Urban Air Mobility
UAS	Uncrewed Aerial Systems
UASOO	UAS Operations Office
UASP	Unmanned Aerial System Pilot Kit
VTOL	Vertical Takeoff and Landing
YOLO	You Only Look Once
ZIA	Zero Impact Aviation
ZIA-HH	Zero Impact Aviation – Human Health

References

1. K. D. Wichman, D. R. Fitzgerald, K. J. Skaggs, K. V. Papathakis, T. Hoang, and T. M. Stinchfield, "Convergent aeronautics solutions project," Presentation at Convergent Aeronautics Solutions (CAS) Project, 2025, https://ntrs.nasa.gov/citations/20240015773.

2. "9-1-1 statistics," https://www.nena.org/page/911Statistics.

3. R. Sampson, "Misuse and abuse of 911," https://popcenter.asu.edu/content/misuse-and-abuse-911-0.

4. H. K. Mell, S. N. Mumma, B. Hiestand, B. G. Carr, T. Holland, and J. Stopyra, "Emergency medical services response times in rural, suburban, and urban areas," *JAMA surgery*, 2017, https://pmc.ncbi.nlm.nih.gov/articles/PMC5831456/.

5. S. King, S. Major, and M. McCollum, "Drone as first responder programs: A new paradigm in policing," 2023, https://www.mitre.org/news-insights/publication/drone-first-responder-programs-new-paradigm-policing.

6. P. King, S. Major, and M. McCollum, "Drone program," https://www.chulavistaca.gov/departments/police-department/programs/uas-drone-program.

7. J. Fody, S. Lehman, and J. T. Slagel, "Uncrewed aerial systems for emergency medical first response: A market research report," NASA, Technical Memorandum, 2024, https://ntrs.nasa.gov/citations/20240011617.

8. P. Cohn, A. Green, M. Langstaff, and M. Roller, "Commercial drones are here: The future of unmanned aerial systems," *McKinsey & Company*, 2017, https://www.mckinsey.com/industries/logistics/our-insights/commercial-drones-are-here-the-future-of-unmanned-aerial-syste.

9. My Drone Services, "Definitive guide to drone use by public safety agencies," 2024. [Online]. Available: https://mydroneservices.com/definitive-guide-to-drone-use-by-public-safety-agencies/

10. A. M. Johnson, C. J. Cunningham, E. Arnold, W. D. Rosamond, and J. K. Zègre-Hemsey, "Impact of using drones in emergency medicine: What does the future hold?" *Open Access Emergency Medicine*, 2021, https://pubmed.ncbi.nlm.nih.gov/34815722/.

11. S. Sanz-Martos, M. D. López-Franco, C. Álvarez-García, N. Granero-Moya, J. M. López-Hens, S. Cámara-Anguita, P. L. Pancorbo-Hidalgo, and I. M. Comino-Sanz, "Drone applications for emergency and urgent care: a systematic review," *Prehospital and Disaster Medicine*, 2022, https://pubmed.ncbi.nlm.nih.gov/35678389/.

12. U.S. Congress, "S.1939 - Federal Aviation Administration Reauthorization Act," 2024, https://www.congress.gov/bill/118th-congress/senate-bill/1939/text.

13. NIST, "2024 first responder UAS open innovation prize challenge," https://www.nist.gov/ctl/pscr/open-innovation-prize-challenges/past-prize-challenges/2024-first-responder-uas, 2024.

14. DARPA, "DARPA triage challenge," 2023, https://www.darpa.mil/research/challenges/darpa-triage-challenge.

15. R. Hawkins, C. Paterson, C. Picardi, Y. Jia, R. Calinescu, and I. Habli, "Guidance on the assurance of machine learning in autonomous systems (AMLAS)," *arXiv preprint*, 2021, https://arxiv.org/abs/2102.01564.

16. K. Dmitriev, J. Schumann, and F. Holzapfel, "Toward certification of machine-learning systems for low criticality airborne applications," in *IEEE/AIAA 40th Digital Avionics Systems Conference, DASC*. IEEE, 2021, https://arxiv.org/abs/2209.13975.

17. Unity Technologies, "Unity," 2023, https://unity.com/.

18. T. Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Computer Vision–ECCV: 13th European conference*, 2014, https://arxiv.org/abs/1405.0312.

19. A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research*, 2013, https://www.cvlibs.net/publications/Geiger2013IJRR.pdf.

20. J. Brown, "CERTAIN: City environment range testing for autonomous integrated navigation," NASA, Technical Memorandum, 2016, https://ntrs.nasa.gov/citations/20160009254.

21. N. Brown, "AIRVUE (crowdsourcing) airborne instrumentation for real-world video of urban environments," in *Transformational Tools and Technologies Project*, 2023, https://ntrs.nasa.gov/citations/20230010665.

22. NASA, "Transformational tools and technologies (TTT) project," https://www.nasa.gov/directorates/armd/tacp/ttt/.

23. S. E. Ebadi, S. Dhakad, S. K. Vishwakarma, C. Wang, Y. C. Jhang, M. Chociej, A. Crespi, A. Thaman, and S. Ganguly, "PSP-HDRI+: A synthetic dataset generator for pre-training of human-centric computer vision models," in *International Conference on Machine Learning*, 2022, https://arxiv.org/abs/2207.05025.

24. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, 2016, https://arxiv.org/abs/1506.02640.

25. M. Sohan, T. Sai Ram, and C. V. Rami Reddy, “A review on YOLOv8 and its advancements,” in *International Conference on Data Intelligence and Cognitive Informatics.* Springer, 2024, https://link.springer.com/chapter/10.1007/978-981-99-7962-2_39.

26. G. Jocher, A. Chaurasia, and J. Qiu, “Ultralytics YOLOv8,” 2023, https://github.com/ultralytics/ultralytics.

27. G. Jocher and J. Qiu, “Ultralytics YOLOv11,” 2024, https://github.com/ultralytics/ultralytics.

28. “AIRVUE techport dataset,” https://techport.nasa.gov/projects/147022.

29. “WESCAM MX-10, air surveillance and reconnaissance,” https://www.l3harris.com/all-capabilities/wescam-mx-10-air-surveillance-and-reconnaissance.

30. “Label studio,” https://labelstud.io/.

31. “NASAspark,” https://www.nasa.gov/directorates/stmd/prizes-challenges-crowdsourcing-program/center-of-excellence-for-collaborative-innovation-coeci/nasa-spark/.

32. B. Fei, J. Xu, R. Zhang, Q. Zhou, W. Yang, and Y. He, “3D Gaussian splatting as new era: A survey,” *IEEE Transactions on Visualization and Computer Graphics*, 2024, https://arxiv.org/abs/2402.07181.

33. J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel, “Domain randomization for transferring deep neural networks from simulation to the real world,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2017, https://arxiv.org/abs/1703.06907.

34. A. Koubâa, *Robot Operating System (ROS).* Springer, 2017, vol. 1, https://doi.org/10.1007/978-3-319-26054-9.

35. P. Huang, S. Ouyang, H. Yan, X. Wang, J. J. Lee, and Q. Zeng, “Effect of emergency medical service response time on fatality risk of freeway crashes: Bayesian random parameters spatial logistic approach,” *Frontiers in Public Health*, 2024, https://pmc.ncbi.nlm.nih.gov/articles/PMC11540930/.

36. Virginia Department of Transportation, “Virginia department of transportation traffic cameras,” https://511.vdot.virginia.gov.

37. US Department of Health and Human Services, “Start adult triage algorithm,” Radiation Emergency Medical Management: REMM, https://remm.hhs.gov/startadult.htm.

38. S. Owre, J. M. Rushby, and N. Shankar, “PVS: A prototype verification system,” in *International Conference on Automated Deduction.* Springer, 1992, https://link.springer.com/chapter/10.1007/3-540-55602-8_217.

39. Computer Science Laboratory, SRI International, "PVS specification language and theorem prover," 2025, https://pvs.csl.sri.com/.

40. M. Clark, X. Koutsoukos, R. Kumar, I. Lee, G. Pappas, L. Pike, J. Porter, and O. Sokolsky, "A study on run time assurance for complex cyber physical systems," Air Force Research Lab, Tech. Rep., 2013, https://leepike.github.io/pubs/RTA-CPS.pdf.

41. A. Goodloe, "Challenges in high-assurance runtime verification," in *International Symposium on Leveraging Applications of Formal Methods*. Springer, 2016, https://ntrs.nasa.gov/api/citations/20160012454/downloads/20160012454.pdf.

42. I. Lee, S. Kannan, M. Kim, O. Sokolsky, and M. Viswanathan, "Runtime assurance based on formal specifications," in *International Conference on Parallel and Distributed Processing Techniques and Applications, PDPTA*, 1999, https://repository.upenn.edu/bitstreams/1cd4ce89-d925-4fa5-9fc2-6d2d6681abf6/download.

43. J. G. Fuller, "Run-time assurance: A rising technology," in *2020 AIAA/IEEE 39th Digital Avionics Systems Conference, DASC*. IEEE, 2020, https://ieeexplore.ieee.org/document/9256425.

44. J. D. Schierman, M. D. DeVore, N. D. Richards, N. Gandhi, J. K. Cooper, K. R. Horneman, S. Stoller, and S. Smolka, "Runtime assurance framework development for highly adaptive flight control systems," 2015, https://apps.dtic.mil/sti/citations/AD1010277.

45. M. Bodson, J. Lehoczky, R. Rajkumar, L. Sha, and J. Stephan, "Analytic redundancy for software fault-tolerance in hard real-time systems," in *Foundations of Dependable Computing: Paradigms for Dependable Applications*. Springer, 1994, https://link.springer.com/chapter/10.1007/978-0-585-27316-7_7.

46. L. Sha, "Using simplicity to control complexity," *IEEE Software*, 2001, https://ieeexplore.ieee.org/document/936213.

47. D. Seto, B. Krogh, L. Sha, and A. Chutinan, "The simplex architecture for safe online control system upgrades," in *Proceedings of the 1998 American Control Conference, ACC*. IEEE, 1998, https://experts.illinois.edu/en/publications/the-simplex-architecture-for-safe-on-line-control-system-upgrades/.

48. J. Yang, M. A. Islam, A. Murthy, S. A. Smolka, and S. D. Stoller, "A simplex architecture for hybrid systems using barrier certificates," in *Computer Safety, Reliability, and Security: 36th International Conference, SAFECOMP*. Springer, 2017, https://www3.cs.stonybrook.edu/~stoller/safecomp2017.html.

49. S. Prajna and A. Jadbabaie, "Safety verification of hybrid systems using barrier certificates," in *International Workshop on Hybrid Systems: Computation and Control*. Springer, 2004, https://link.springer.com/chapter/10.1007/978-3-540-24743-2_32.

50. T. T. Johnson, S. Bak, M. Caccamo, and L. Sha, "Real-time reachability for verified simplex design," *ACM Transactions on Embedded Computing Systems, TECS*, 2016, https://ieeexplore.ieee.org/document/7010482.

51. D. T. Phan, R. Grosu, N. Jansen, N. Paoletti, S. A. Smolka, and S. D. Stoller, "Neural simplex architecture," in *NASA Formal Methods: 12th International Symposium, NFM.* Springer, 2020, https://arxiv.org/abs/1908.00528.

52. G. Brat and G. Pai, "Runtime assurance of aeronautical products: Preliminary recommendations," NASA, Technical Memorandum, 2023, https://ntrs.nasa.gov/api/citations/20220015734/downloads/tm-rta-guidance.pdf.

53. ASTM International, "Standard practice for methods to safely bound behavior of aircraft systems containing complex functions using runtime assurance," ASTM International, 2021, https://store.astm.org/f3269-21.html.

54. L. V. Jospin, H. Laga, F. Boussaid, W. Buntine, and M. Bennamoun, "Hands-on Bayesian neural networks—a tutorial for deep learning users," *IEEE Computational Intelligence Magazine*, 2022, https://arxiv.org/abs/2007.06823.

55. E. Goan and C. Fookes, "Bayesian neural networks: An introduction and survey," in *Case Studies in Applied Bayesian Data Science.* Springer, 2020, https://arxiv.org/abs/2006.12024.

56. S. Owre, N. Shankar, J. M. Rushby, and D. W. Stringer-Calvert, "PVS language reference," *Computer Science Laboratory, SRI International*, 1999, https://pvs.csl.sri.com/doc/pvs-language-reference.pdf.

57. J. Siratt, "Interactive theorem proving and applications at NASA," Presentation at Formal Methods Cross-Lab Meeting, Johns Hopkins University Applied Physics Laboratory, 2025, https://ntrs.nasa.gov/api/citations/20250001586/downloads/siratt_presentation.pdf.

58. C. A. Muñoz, A. Dutle, A. Narkawicz, and J. Upchurch, "Unmanned aircraft systems in the national airspace system: a formal methods perspective," *ACM SIGLOG News*, 2016, https://shemesh.larc.nasa.gov/fm/papers/SIGLOG-MDNU-draft-2016.pdf.

59. A. Dutle, M. Moscato, L. Titolo, C. Munoz, G. Anderson, and F. Bobot, "Formal analysis of the compact position reporting algorithm," *Formal Aspects of Computing*, 2021, https://shemesh.larc.nasa.gov/fm/papers/VSTTE2017-draft.pdf.

60. J. T. Slagel, L. M. White, A. Dutle, C. A. Muñoz, and N. Crespo, "A verification framework for runtime assurance of autonomous UAS," in *43rd Digital Avionics Systems Conference, DASC.* IEEE, 2024, https://ieeexplore.ieee.org/document/10748654.

61. C. Urban and A. Miné, "A review of formal methods applied to machine learning," 2021, https://arxiv.org/abs/2104.02466.

62. A. Dutle, J. T. Slagel, and J. V. Siratt, “Verification of autonomous systems,” 2025, https://ntrs.nasa.gov/citations/20250003529.

63. J. Redmon and A. Farhadi, “YOLO9000: better, faster, stronger,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, https://arxiv.org/abs/1612.08242.

64. S. Tobin-Hochstadt, V. St-Amour, E. Dobson, and A. Takikawa, “The typed racket reference,” 2014, https://docs.racket-lang.org/ts-reference.

65. J. T. Slagel, M. Moscato, L. White, C. A. Muñoz, S. Balachandran, and A. Dutle, “Embedding differential dynamic logic in PVS,” in Proceedings 18th International Workshop on *Logical and Semantic Frameworks, with Applications*. Open Publishing Association, 2024, https://shemesh.larc.nasa.gov/fm/papers/LSFA_23_submit.pdf.