Heads-Up, Hands-Off Medic-Robot Teaming with Foundation Models and Agentic AI

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

11100 Johns Hopkins Road
Laurel, MD 20723-6099

Johns Hopkins Applied Physics Laboratory
Research & Exploratory Development Mission Area – Frontier Intelligent Systems
Global Health Mission Area – Assured Care

Collaboration with Army Research Laboratory
U.S. Army Combat Capabilities Development Command
Army Artificial Intelligence Innovation Institute

David Handelman, Ph.D.
Principal Research Scientist
Robotics & Autonomous Systems
David.Handelman@jhuapl.edu

5 August 2026

Disclosure

- David Handelman, Ph.D. has no conflicts to report.

- This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-21-2-0211. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. This research was also supported by Johns Hopkins Applied Physics Laboratory Independent Research and Development funding. DISTRIBUTION A. Approved for public release; distribution unlimited

Research Aim: Extend Medic Capacity Through Human-Robot Teaming

- **Motivation – Sustain care under pressure**
  - Help medics manage more casualties in dangerous, resource-constrained environments without navigating complex robotic interfaces.
- **Research Question – Can robots become effective teammates?**
  - Can foundation models and agentic AI enable natural interaction, coordinated action, and shared understanding while preserving medic authority and system trustworthiness?
- **Technical Objective – Advance toward medical assistance**
  - Integrate speech, multimodal perception, reasoning, and edge autonomy – progressing from conversational teaming and scene assessment to supervised medical task execution.

Approach: Intent-Based Autonomy for Medic-Robot Teaming

- **Hypothesis – Team through shared intent**
  - Foundation models and agentic AI can help medics direct heterogeneous robots through natural interaction rather than low-level control.
- **Approach – Progressively increase autonomy**
  - Shift between teleoperation, shared control, and high-level tasking while preserving medic authority and reducing interface burden.
- **Architecture – Integrate generative, neuro-symbolic, and agentic AI**
  - Leverage generative AI for natural interaction through speech (LLM) and robot perception (VLM), and agentic techniques for task planning, execution, monitoring, and replanning.
  - Neuro-symbolic methods combine the rigor and constraint awareness of formal planning methods with the creativity and adaptability of large language models.
- **Teaming progression:** Build from robot maneuver and information exchange toward collaborative decision support and medical task execution

The Problem

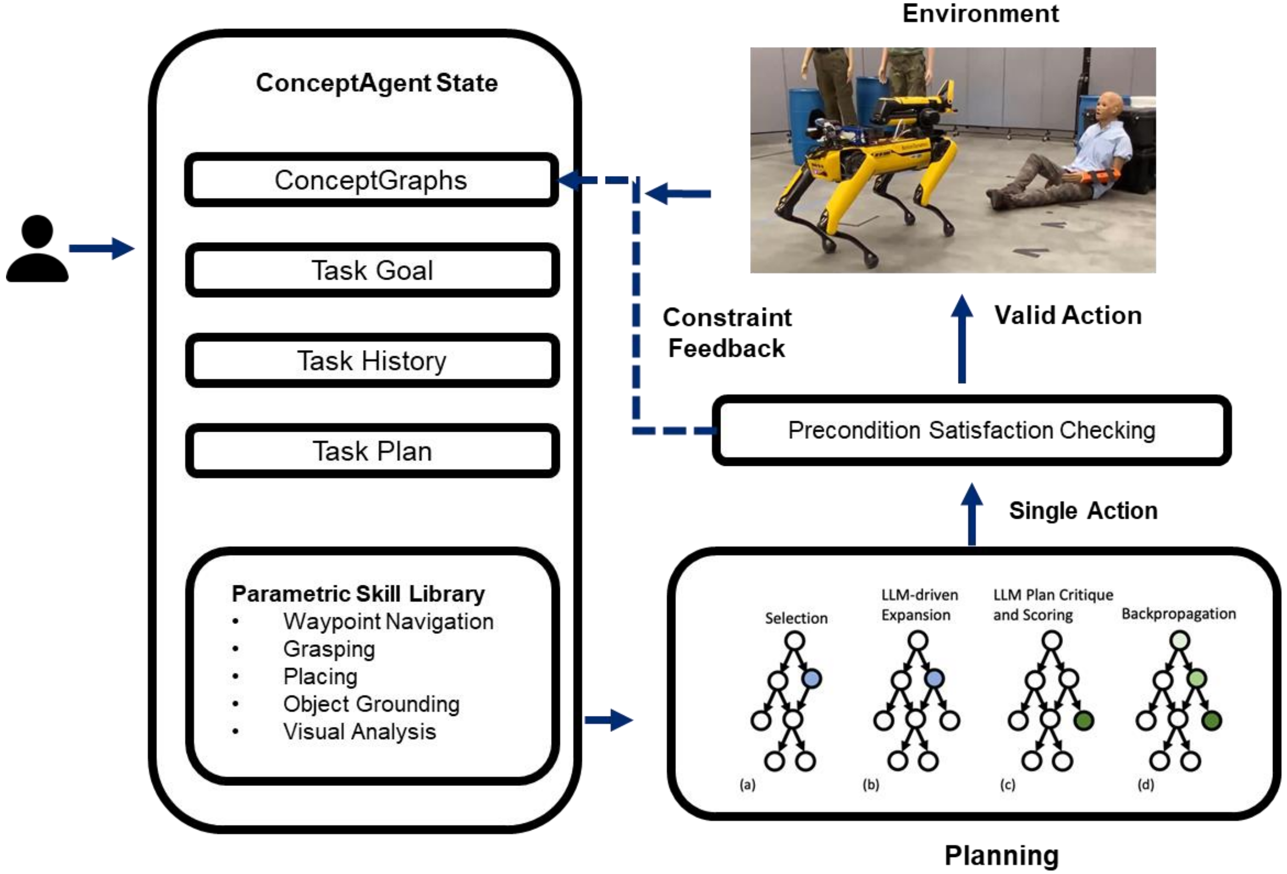
"We've got to get to a point where the Soldier never lowers his or her weapon to pick up a remote control. That's the golden rule. That's my red line."

Brigadier General Phil Kiniery
U.S. Army Maneuver Center of Excellence
AUSA MCoE Industry Day, April 2025

Agentic Robot AI for Medic-Robot Teaming

Desired Capability

Give Command	Natural Language
Understand the Surroundings	**Robot** creates 3D world map
Create an Execution Plan	**Robot** creates execution plan and subgoals
Perform Tasks	**Robot** executes
Evaluate Progress and Replan as Needed	**Robot** evaluates and replans



Byrd et al. Constrained natural language action planning for resilient embodied systems. npj Artificial Intelligence 2026
Rivera et al. ConceptAgent: LLM-driven precondition grounding and tree search for robust task planning and execution. ICRA 2025
Gu et al. ConceptGraphs: Open-vocabulary 3d scene graphs for perception and planning. ICRA 2024
Handelman et al. Leveraging foundation models for scene understanding in human-robot teaming. SPIE 2024
Handelman et al. Multi-agent playbook for human-robot teaming. SPIE 2023

Autonomy and Teaming Software Stack

APL Intelligent Mission System (IMS)

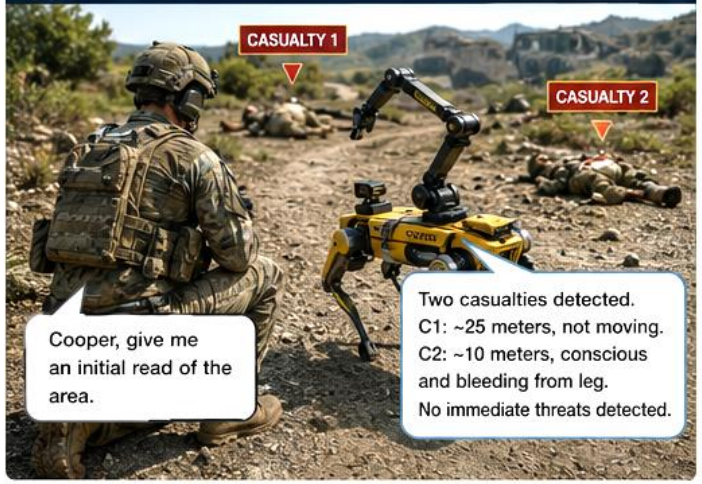
Mission applications	**EMS support · search and rescue · disaster response · mass casualty triage · combat casualty care**
Human-robot teaming	**Voice/radio · gesture · shared situational awareness · mixed reality**
Autonomy engine	**Perception · navigation · planning · task execution · team coordination**
Edge compute payload	**GPU compute · sensor fusion · comms integration**
COTS robot hardware	**Drones · UGVs · robot dogs · vendor platforms**

- Hardware-agnostic software and payload integration for drones, UGVs, quadrupeds, and future robotic platforms
- Enable first responders to communicate intent, have better situational awareness, and delegate useful actions to robots

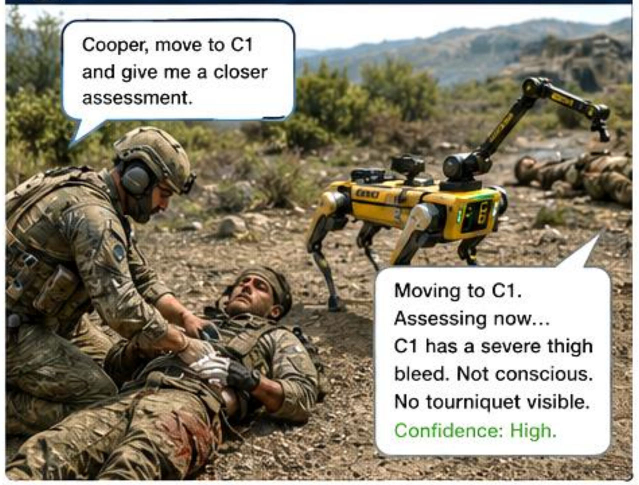
SCENARIO: PARALLEL TRIAGE AND TREATMENT WITH ROBOTIC TEAMMATE

One medic. Two casualities. Heads-up, hands-off teaming.

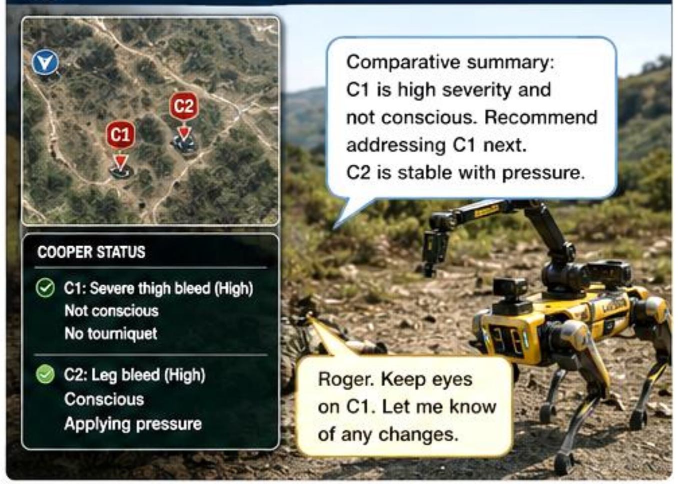
1 ARRIVAL & INITIAL ASSESSMENT 00:00–00:25



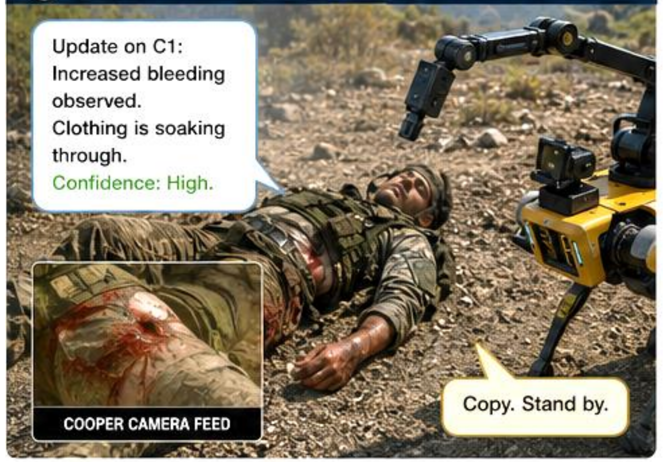
2 DELEGATED ASSESSMENT 00:25–01:05



3 PRIORITIZATION & CONTEXT 01:05–01:35

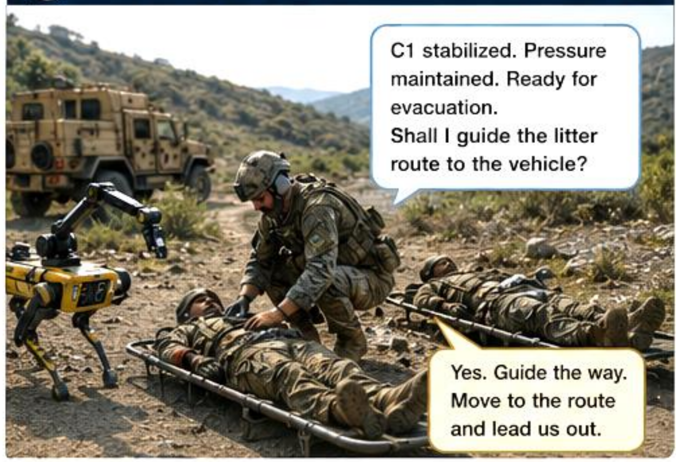


4 OVERWATCH & CHANGE DETECTION 01:35–02:15



5 ASSISTED INTERVENTION (HANDS SUPPORT) 02:15–03:00

6 HANDOFF & EVAC PREP 03:00–03:30



THE TAKEAWAY: The robot extends the medic's eyes and reach—enabling parallel assessment, continuous monitoring, and informed decisions while preserving medic authority and focus on care.

SECOND SET OF EYES

THIRD ARM (EMERGING)

MEDIC IN COMMAND


MP4


165 lb Dummy - Remote Control

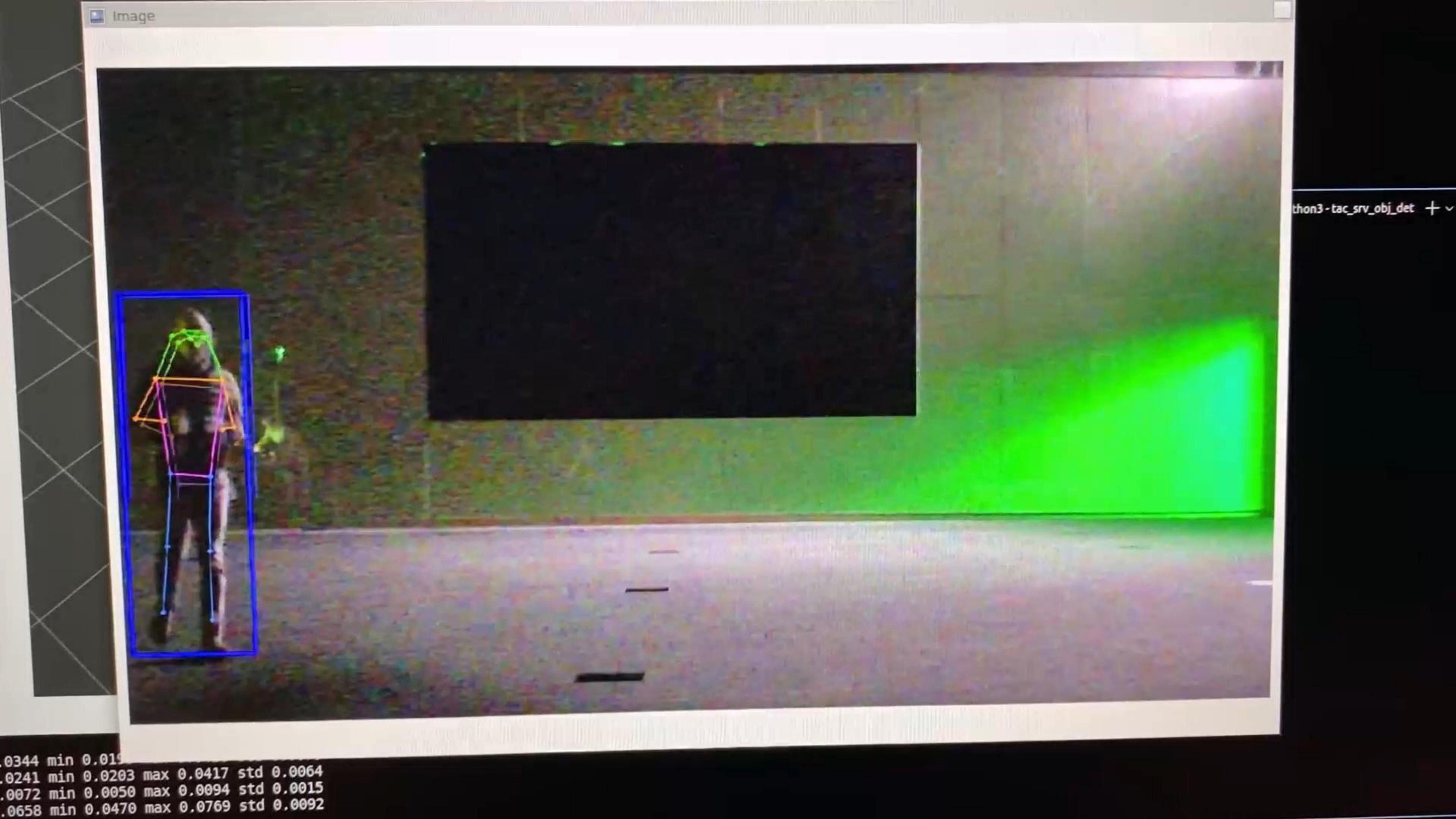
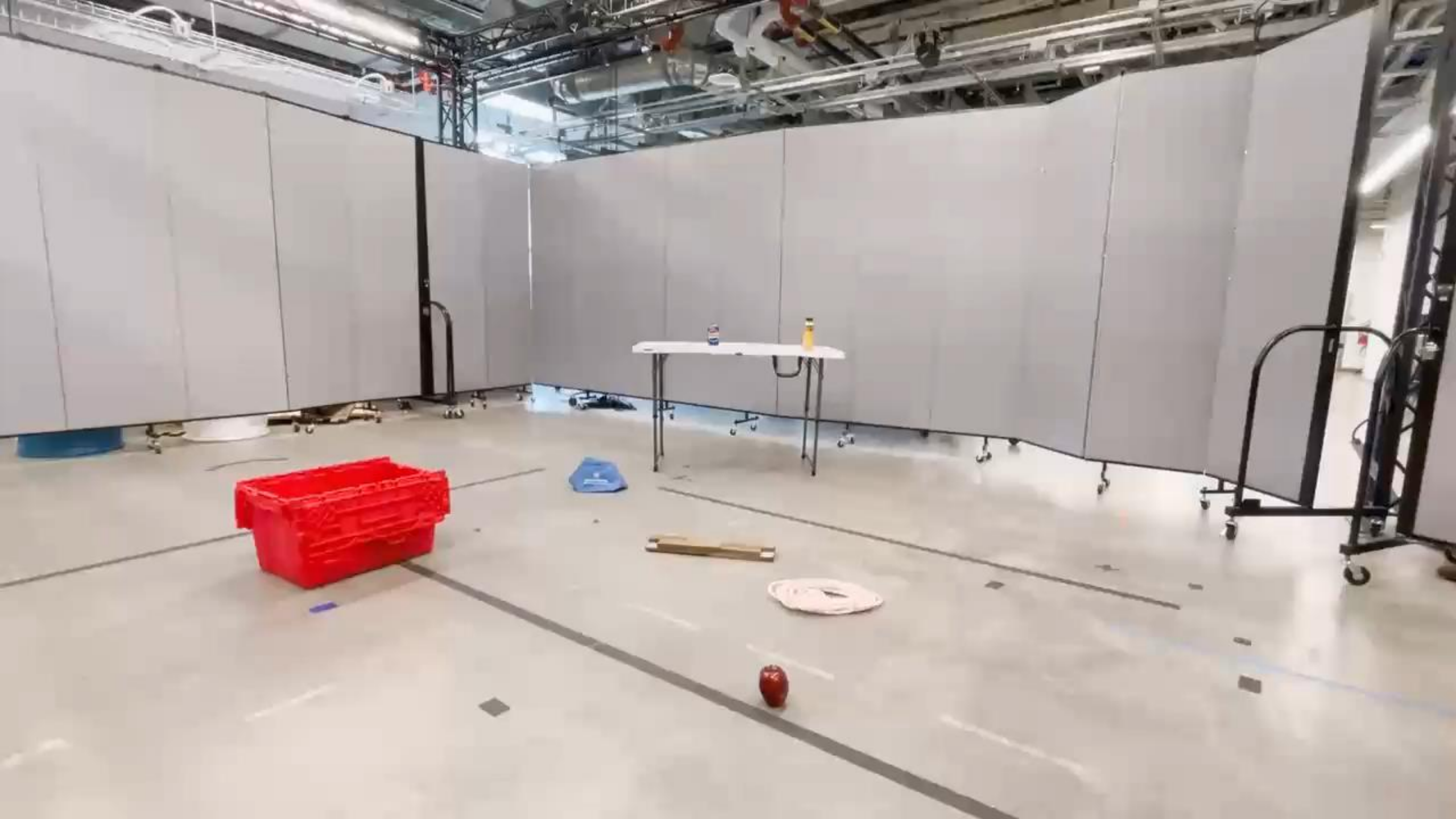
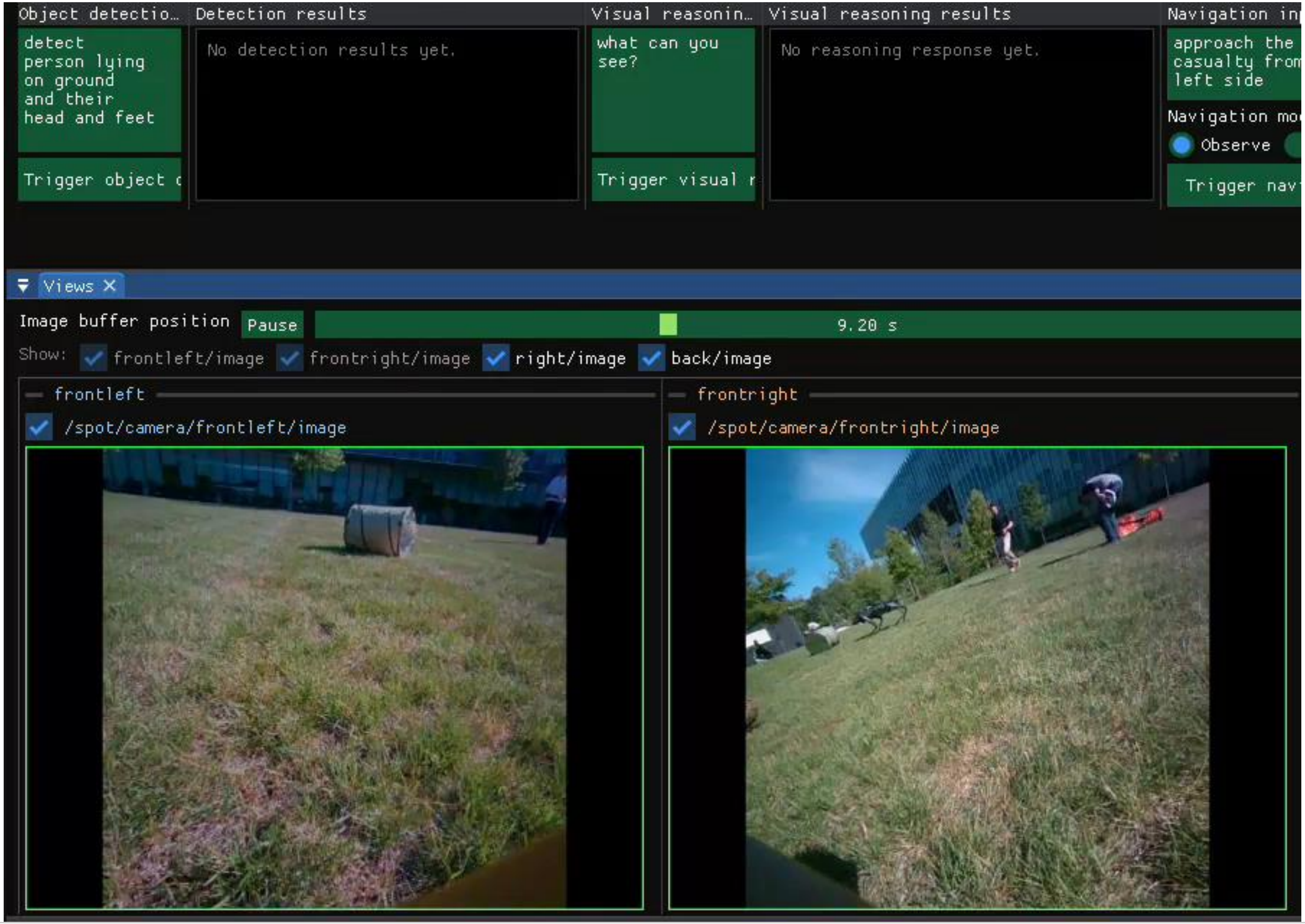


Image
thon3 - tac_srv_obj_det
.0344 min 0.019
.0241 min 0.0203 max 0.0417 std 0.0064
.0072 min 0.0050 max 0.0094 std 0.0015
.0658 min 0.0470 max 0.0769 std 0.0092

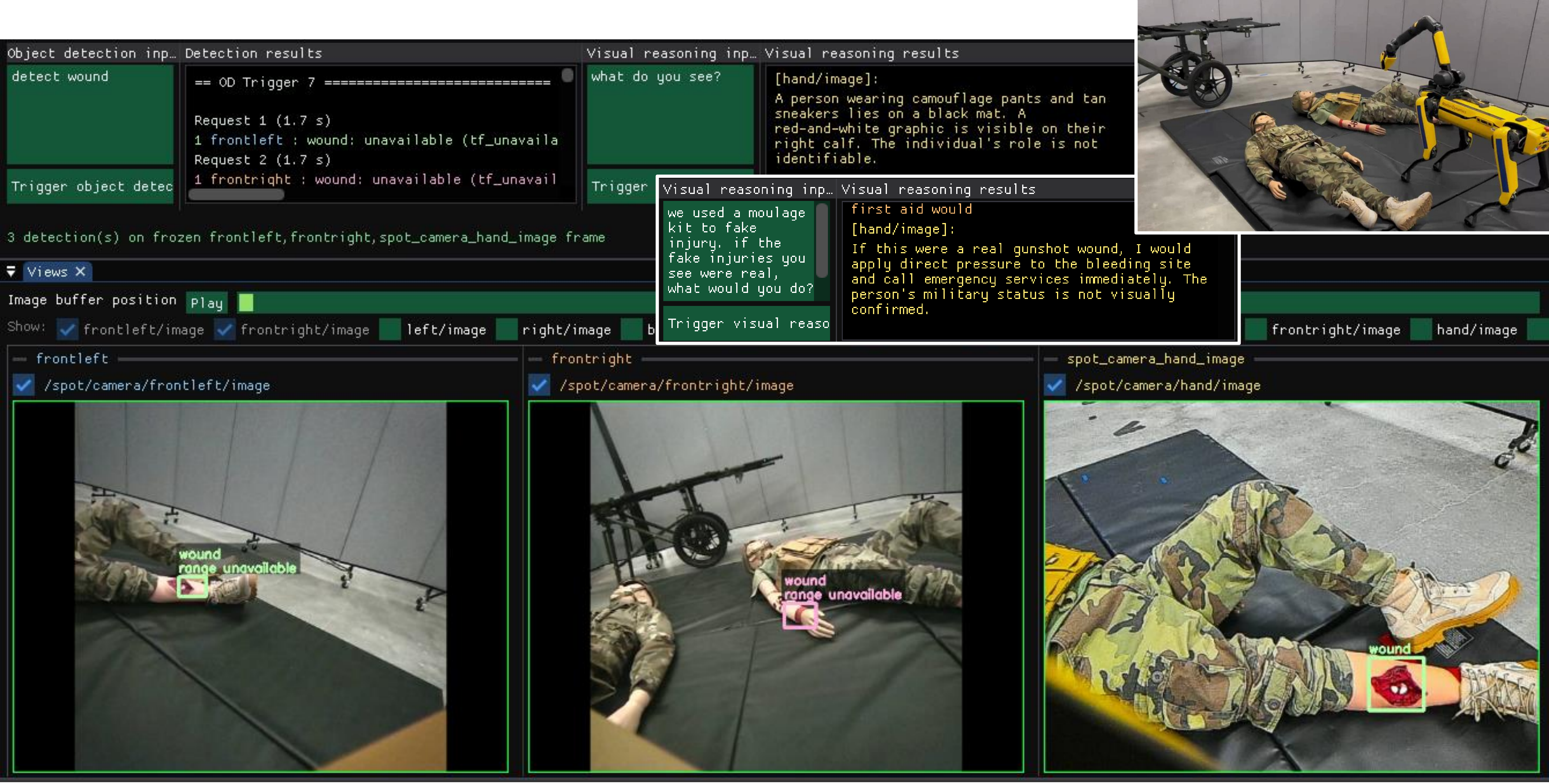



Boston Dynamics

Edge Foundation Model for Object Detection and Visual Reasoning



- Gemma 4:12B – encoder-free, decoder-only open multimodal model natively processes text, vision, and audio via direct linear projection layers
- Optimized for edge hardware (E2B, E4B, 12B, 26B, 31B)
- Open-vocabulary object detection locates specific items using natural language prompts, returning structured bounding box coordinates without custom training
- Visual reasoning and question answering
- Agentic workflows: multi-step reasoning and tool calling.
- Direct application to robot autonomy for casualty care


Object detection inp…
detect wound
Trigger object detec
Detection results
== OD Trigger 7 ============================
Request 1 (1.7 s)
1 frontleft : wound: unavailable (tf_unavaila
Request 2 (1.7 s)
1 frontright : wound: unavailable (tf_unavail
Visual reasoning inp…
what do you see?
Trigger
Visual reasoning results
[hand/image]:
A person wearing camouflage pants and tan sneakers lies on a black mat. A red-and-white graphic is visible on their right calf. The individual's role is not identifiable.
Visual reasoning inp…
we used a moulage kit to fake injury. if the fake injuries you see were real, what would you do?
Trigger visual reaso
Visual reasoning results
first aid would
[hand/image]:
If this were a real gunshot wound, I would apply direct pressure to the bleeding site and call emergency services immediately. The person's military status is not visually confirmed.
3 detection(s) on frozen frontleft,frontright,spot_camera_hand_image frame
Views
Image buffer position
Play
Show:
frontleft/image
frontright/image
left/image
right/image
frontright/image
hand/image
frontleft
/spot/camera/frontleft/image
frontright
/spot/camera/frontright/image
spot_camera_hand_image
/spot/camera/hand/image
wound
range unavailable
wound
range unavailable
wound

Edge Foundation Model for Visual Reasoning – Soldier Rank Identification

Preliminary Results: Enabling Capabilities and Path to Medical Support

- **Demonstrated autonomy building blocks**: IMS prototypes highlight limited natural-language interaction, coordinated maneuver, and shared situational awareness within human-robot teams.
- **Intent-based teaming**: Agentic methods like ConceptAgent ground operator intent in perceived objects, spatial concepts, and robot capabilities – supporting dialogue beyond predefined commands.
- **Agentic planning and coordination**: Neurosymbolic techniques such as FLEET and SCLPlan combine formal planning methods with the creativity and adaptability of large language models, enabling goal-directed planning and heterogeneous robot coordination, but are currently low-TRL.
- **Transitioning toward medical assistance**: We are integrating these capabilities in manikin-based casualty scenarios, progressing from scene assessment and information sharing toward supervised task assistance and medic-in-the-loop evaluation.

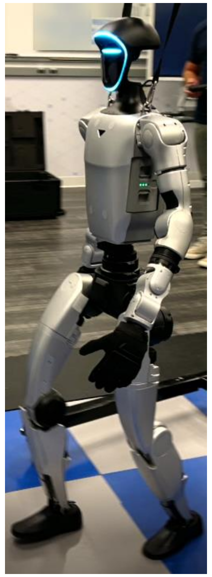
Future Directions / APL Projects – Combat Casualty Care Scenario

- **Adaptive Human-Robot Teaming** – ARL CA and APL IRAD (prior)
- **Medic-Robot Teaming for Collaborative Intervention** – APL IRAD (prior)
- **Autonomous Patient Evacuation Robot** – APL IRAD (prior)
- **Shared Situational Awareness for Distributed Operations** – ARL CA
  ARL CHAMP multi-agent SLAM, Soldier Borne Mission Command (SBMC): Tactical XR – fuses data from Soldiers, sensors, and mission systems into one clear operational picture, day or night
- **Open-World Reasoning and Decision Making** – ARL CA
  Flying agentic drones at ARL R2C2, DARPA ITM-inspired CCC scenario – no good options, rules of engagement, commanders intent, constitutional AI
- **Collaborative Autonomous Robotic Engagement** – DARPA Seedling
- **Robotic Manipulation Autonomy using VLAs** – APL IRAD
  Quadruped learning from demonstration/imitation learning
- **Other agentic UGV and UAS efforts**
  Drone control with Soldier-borne compute, autonomous robotic ISR, site security

Search Intubation Bagging

Formation following and shared situational awareness

Casualty evacuation



Acknowledgements

Sponsors

- Army Research Laboratory
- DARPA BTO
- APL Research and Exploratory Development Mission Area
- APL Global Health Mission Area

APL Team


David Handelman, Ph.D.


Corban Rivera, Ph.D.


Christopher Korpela, Ph.D.


Emma Holmes


Bethany Kemp


Eric Bubar, Ph.D.


Allie Gaines, Ph.D.


Meghan Booker, Ph.D.


Grayson Byrd


THIS WE'LL DEFEND
250
OPTWS

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY