

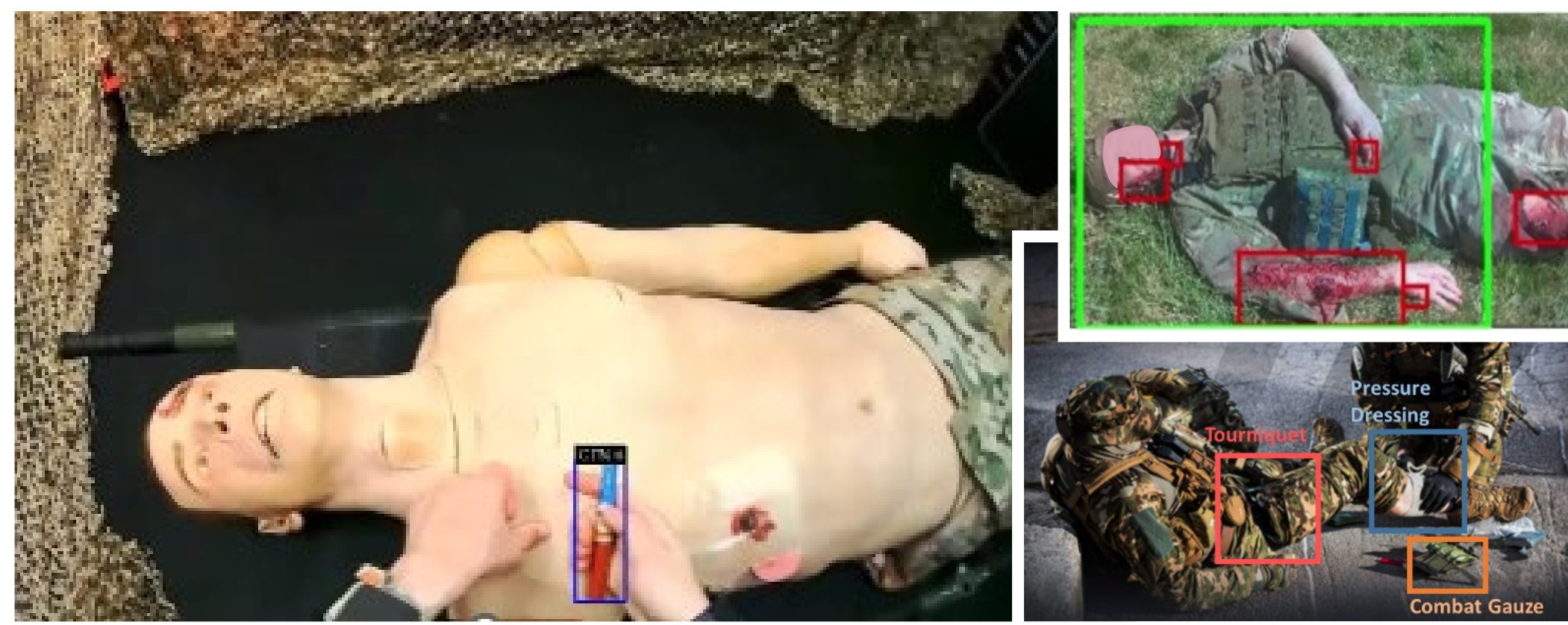
Performance Assessment Methodology and End-to-End Framework for Tactical Combat Casualty Care Autonomous Documentation Algorithms

Jaeyeon Lee, PhD¹, Jeanette R. Little, MS¹, Ethan Quist, ME², Omar Badawi, PharmD¹, MPH, Nathan Fisher, MS¹

¹U.S. Army Institute of Surgical Research, ²Defense Advanced Research Projects Agency

Background

- Tactical Combat Casualty Care (TCCC) requires rapid, high-intensity medical interventions where documentation is often delayed or incomplete [1]
- Human Machine Teaming (HMT) with autonomous documentation improves continuity of care and operational efficiency [2]
- Conventional artificial intelligence (AI) evaluation methods do not capture the workflow-dependent nature of TCCC care



Objective

- Develop a standardized evaluation framework for autonomous TCCC documentation algorithms in realistic field care workflows
- Assess integrated system-level algorithm performance under edge-device resource constraints
- Identify strengths, limitations, and areas requiring model refinement to support iterative performance improvement

Approach

- Evaluated three detection tasks: treatment procedures, medical objects, and injury locations
- Assessed representative algorithms within an orchestration framework for edge-device deployment using detection metrics and latency analysis
- Evaluated system-level computational efficiency using CPU, GPU, memory utilization, and throughput metrics

References:

- Joint Trauma System (JTS), Committee on Tactical Combat Casualty Care (CoTCCC). "Tactical Combat Casualty Care (TCCC) Guidelines for All Service Members" 2024.
- U.S. Army Futures Command (2022) Army Futures Command Concept for Medical 2028, AFC Pamphlet 71-20-12. U.S. Army Medical Center of Excellence.

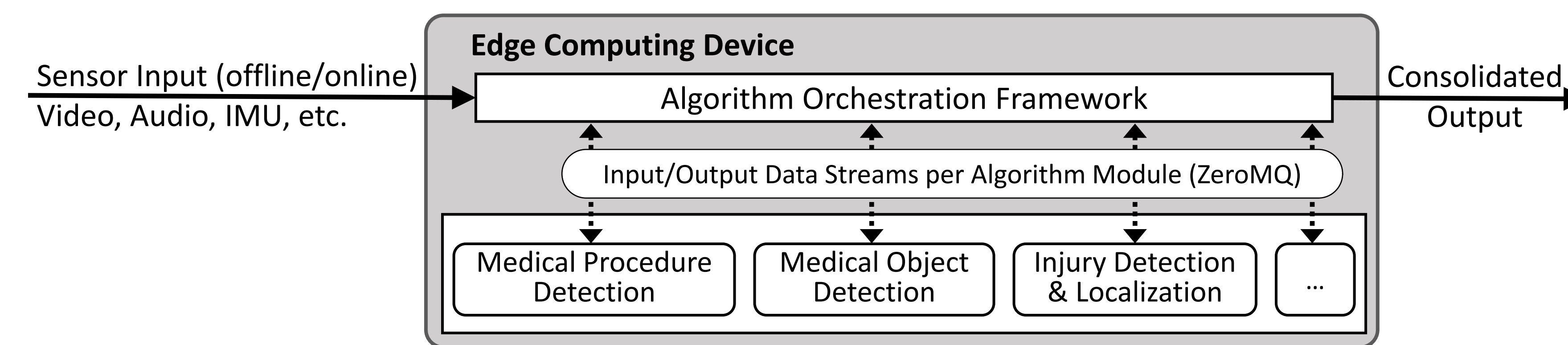
Acknowledgements:

The following individuals are acknowledged for their direct and indirect support to this effort: Dr. Shawn Burgener, Dr. Jonathan Chambers, Dr. Brant Chee, Mr. Peter Chung, Mr. Valmik Desai, Ms. Alix Donnelly, Mr. David Drewry III, Ms. Stephanie Edwards, Mr. Tyler Feldman, Dr. Amanda Galante, Dr. Prudhvi Gurram, Ms. Tina Hamrick, Mr. Wesley Huff, Mr. Jakob Kaminsky, Ms. Buffy Krall, Dr. Yemeserach Mekonnen, LTC (Ret) Rachel Morgans, COL (Ret) Jeremy Pamplin, Mr. Matt Quinn, Mr. Jeff Robbins, Ms. Kate Tallaksen, Ms. Ha Tran, and SFC Justin Warren

Disclaimer: The views, opinions and/or findings contained in this presentation are those of the author and do not necessarily reflect the views of the Department of War and should not be construed as an official DoW/Army position, policy or decision unless so designated by other documentation. No official endorsement should be made. Reference herein to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the U.S. Government.

Methods

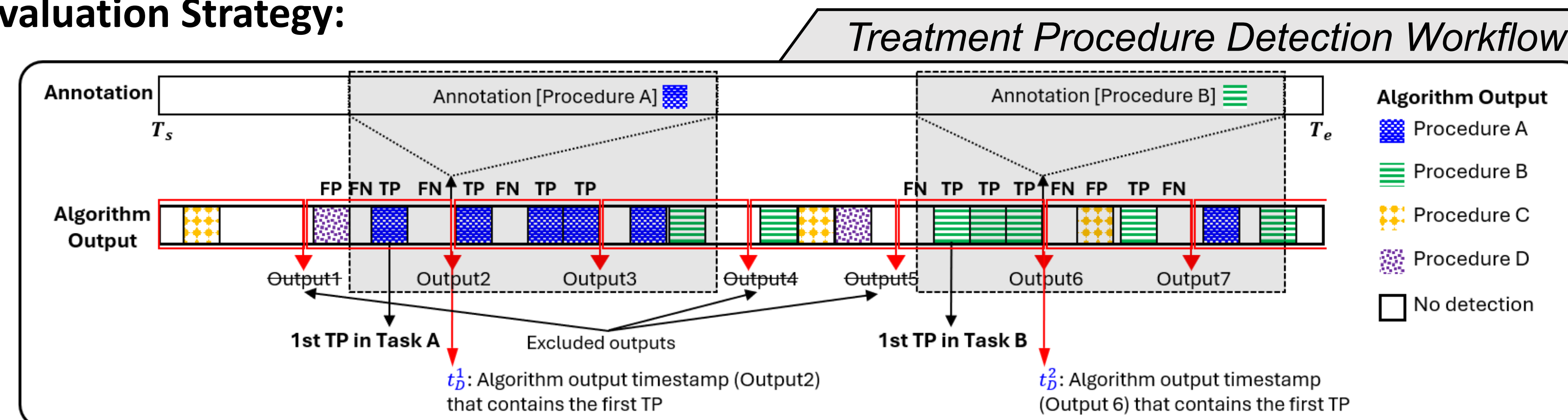
Algorithm Orchestration Process:



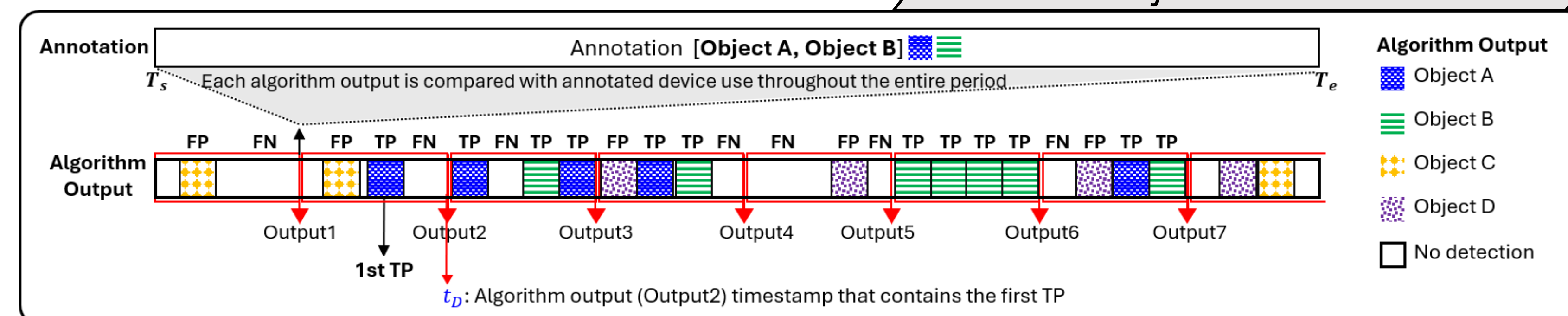
Ground-Truth Annotations:

- Expert combat medics generated event-level TCCC annotations
- Procedure intervals and anatomical regions without frame-level or bounding-box labels

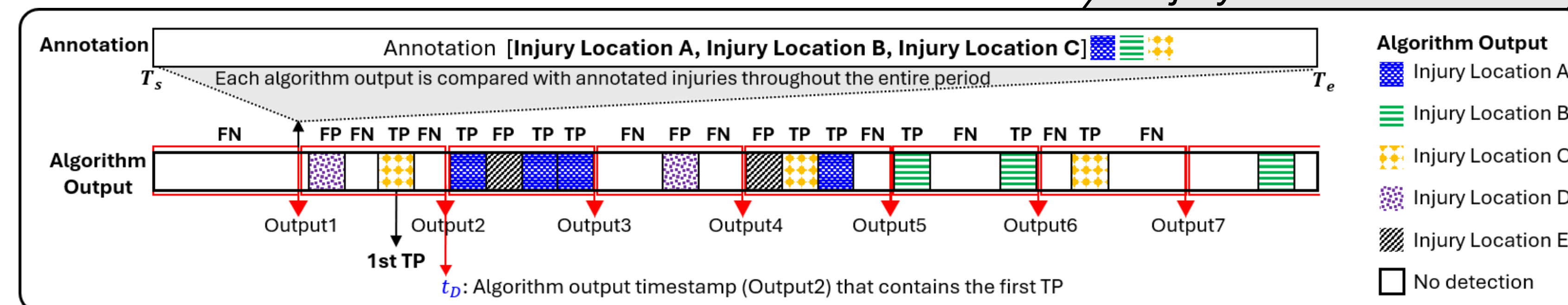
Evaluation Strategy:



Medical Object Detection Workflow



Injury Location Workflow

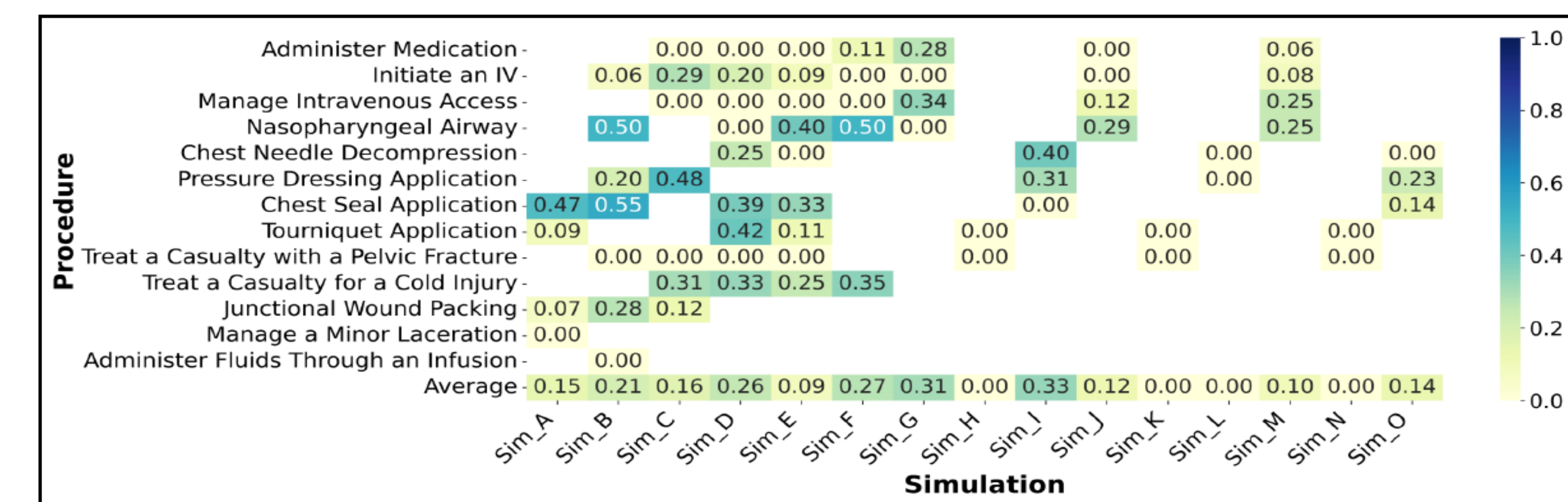
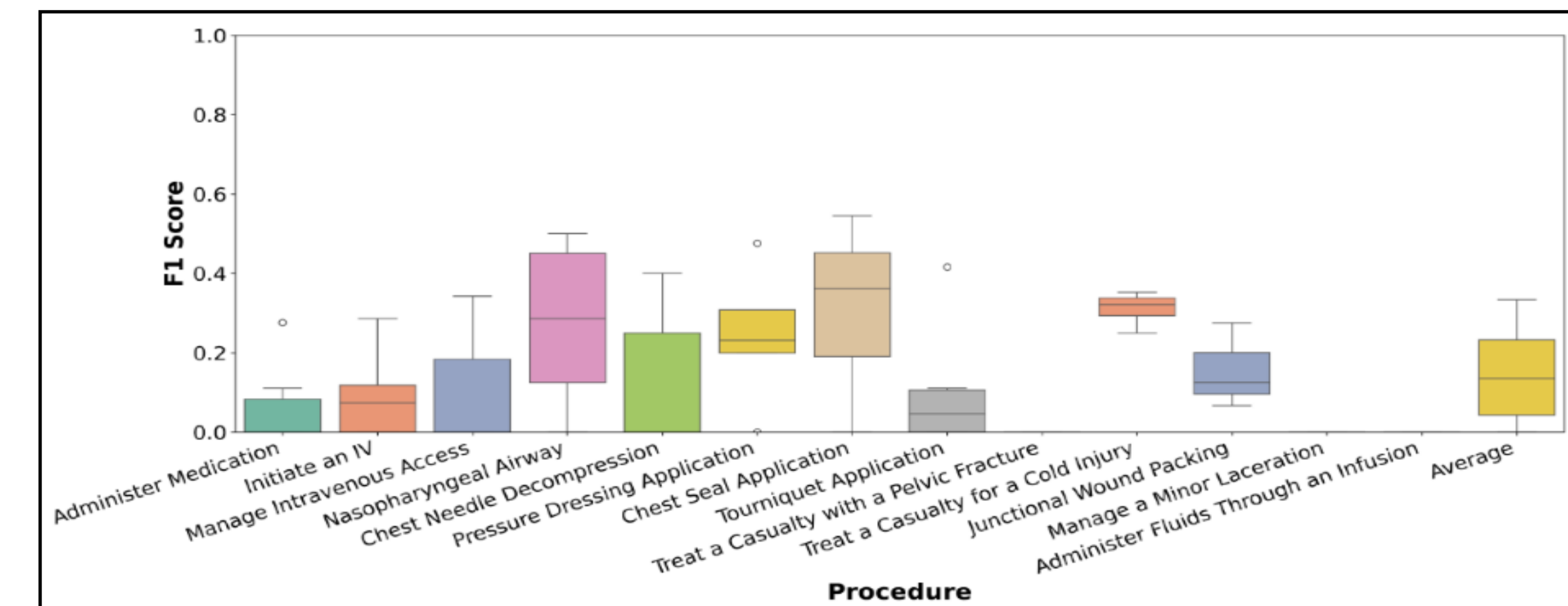


Evaluation Metrics:

Detection Effectiveness		Processing Efficiency	Computational Efficiency	CPU/GPU utilization
Modified accuracy (No TN)	Recall	Lag-time analysis		Memory usage
Precision	F1-score			Throughput

Results

Treatment Procedure Detection:



- Treatment detection varied across procedures and simulations

Medical object detection & injury location:

Simulation	Object Detection				Simulation	Injury Location			
	Accuracy	Precision	Recall	F1		Accuracy	Precision	Recall	F1
Sim_A	0.349	0.406	0.713	0.517	Sim_A	0.143	0.150	0.750	0.250
Sim_B	0.544	0.643	0.780	0.705	Sim_B	0.143	0.151	0.750	0.251
Sim_C	0.247	0.304	0.567	0.396	Sim_C	0.095	0.100	0.867	0.174
Sim_D	0.123	0.112	0.784	0.397	Sim_D	0.186	0.186	0.750	0.314
Sim_E	0.439	0.537	0.707	0.610	Sim_E	0.193	0.202	0.800	0.323
Sim_F	0.149	0.256	0.262	0.259	Sim_F	0.211	0.235	0.667	0.348
Sim_G	0.313	0.398	0.593	0.477	Sim_G	0.000	0.000	0.000	0.000
Sim_H	0.000	0.000	0.000	0.000	Sim_H	0.000	0.000	0.000	0.000
Sim_I	0.419	0.255	0.619	0.391	Sim_I	0.263	0.286	0.697	0.416
Sim_J	0.202	0.236	0.586	0.337	Sim_J	0.429	0.529	0.692	0.600
Sim_K	0.240	0.316	0.500	0.387	Sim_K	0.143	0.150	0.750	0.250
Sim_L	0.156	0.196	0.435	0.270	Sim_L	0.090	0.094	0.852	0.165
Sim_M	0.330	0.412	0.623	0.453	Sim_M	0.302	0.322	0.825	0.463
Sim_N	0.093	0.109	0.387	0.170	Sim_N	0.095	0.100	0.867	0.174
Sim_O	0.209	0.264	0.500	0.346	Sim_O	0.159	0.168	0.749	0.274

Object Detection

Injury Location

- Object detection: higher recall and F1 with variability
- Injury location: recall consistently exceeded precision

Conclusion

- Enabled standardized, workflow-aware evaluation of autonomous TCCC documentation algorithms
- Provided a repeatable framework for objective assessment across diverse TCCC documentation tasks
- Identified algorithm limitations and priorities for future refinement



To learn more, visit:
usa.isr.health.mil