

A Bayesian Reasoning Framework for Robotic Systems in Autonomous Casualty Triage

Szymon Rusiecki^{1,2}, Cecilia Morales², Pia Störy^{2,3},
Kimberly Elenberg², Leonard Weiss⁴, and Artur Dubrawski²

Abstract—Autonomous robots deployed in mass casualty incidents (MCI) face the challenge of making critical decisions based on incomplete and noisy perceptual data. We present an autonomous robotic system for casualty assessment that fuses outputs from multiple vision-based algorithms, estimating signs of severe hemorrhage, visible trauma, or physical alertness, into a coherent triage assessment. At the core of our system is a Bayesian network, constructed from expert-defined rules, which enables probabilistic reasoning about a casualty’s condition even with missing or conflicting sensory inputs. The system, evaluated during the DARPA Triage Challenge (DTC) in realistic MCI scenarios involving 11 and 9 casualties, demonstrated a nearly three-fold improvement in physiological assessment accuracy (from 15% to 42% and 19% to 46%) compared to a vision-only baseline. More importantly, overall triage accuracy increased from 14% to 53%, while the diagnostic coverage of the system expanded from 31% to 95% of cases. These results demonstrate that integrating expert-guided probabilistic reasoning with advanced vision-based sensing can significantly enhance the reliability and decision-making capabilities of autonomous systems in critical real-world applications.

I. INTRODUCTION

When a terrorist attack struck the Boston Marathon in 2013, emergency responders faced around 250 casualties in minutes, overwhelming the city’s trauma treatment capacity and forcing life-or-death triage decisions under extreme pressure [1]. Such mass casualty incidents (MCIs), resulting from terrorist attacks, armed conflicts, or large-scale natural disasters, can involve hundreds of victims and instantly overwhelm local emergency services [2]. In these chaotic scenarios, conventional manual triage methods can fail when faced with high volumes of casualties and complex injury profiles [3], [4].

To address these challenges, autonomous robotic systems offer a compelling solution, capable of surveying disaster zones to identify victims and remotely assess their injuries [5]–[7]. However, deploying these robots in real-world MCI environments introduces a unique set of severe operational challenges [8]–[11]. Robots must navigate cluttered and unpredictable disaster zones, operate under poor visibility from smoke or dust, and contend with severe sensor occlusions where casualties are partially or fully obscured



Fig. 1. UGV and UAV robots used in our experiments

by debris. Consequently, the data they collect with their sensors are often noisy, incomplete, and fragmented, making a reliable assessment difficult.

This inherent unreliability of sensory data poses a fundamental challenge to the perception modules running on these robots. Current perception algorithms often suffer from three critical limitations in this context: **algorithmic isolation**: treating each vital sign independently, **context blindness**: ignoring clinical inter-dependencies, and **brittleness**: underperforming under noisy conditions.

For example, a robot may flag an amputation and a blood detector may see blood, yet without context-aware reasoning the system may not be able to infer that the casualty is actively hemorrhaging, a higher triage priority than an isolated lower-extremity wound. Therefore, autonomous triage requires more than a set of independent detectors; it needs a cognitive architecture that can reason under uncertainty and fuse partial evidence into a coherent assessment.

To overcome these limitations and enable robust robotic decision-making, we propose a cognitive architecture centered on a Bayesian Network (BN). This reasoning engine integrates outputs from multiple perception algorithms into a unified probabilistic estimate of the patient’s condition. Constructed from structured clinical knowledge elicited from experts, our BN is designed to support inference even when a sensor fails or a casualty is occluded, improving the consistency and interpretability of the robot’s triage decisions in real time. This expert-knowledge foundation provides a robust and transparent alternative for decision support in the data-scarce scenarios typical for robotics in MCIs.

In this work, we introduce a knowledge-based cognitive architecture for automated robotic triage. Our key contribu-

¹AGH University of Krakow, Kraków, Poland
rusiecki@student.agh.edu.pl

²Carnegie Mellon University, Pittsburgh, PA, USA
{cgmorale, kelenber, awd}@andrew.cmu.edu

³Osnabrück University, Osnabrück, Germany
pstoery@uni-osnabrueck.de

⁴University of Pittsburgh, Pittsburgh, PA, USA
weissls2@upmc.edu

tions are (1) a novel architecture for autonomous robotic triage centered on an expert-elicited Bayesian Network that integrates multimodal perception inputs, (2) a robust reasoning framework that enables the robot to handle the noisy, uncertain, and incomplete sensory data characteristic of real-world disaster environments, and (3) a full system validation through realistic field scenarios, demonstrating significantly improved triage accuracy over the alternative employing independent algorithmic outputs. The development and evaluation of this system were carried out in the context of the DARPA Triage Challenge (DTC), an initiative aimed at advancing autonomous medical triage capabilities for mass casualty incidents.

To the best of our knowledge, the proposed system represents one of the first neuro-symbolic frameworks for this application to be validated at scale. Its performance was benchmarked in a comprehensive field evaluation involving 20 distinct casualty cases, demonstrating its robustness outside of a controlled laboratory environment.

II. RELATED WORK

A. Bayesian Networks

Bayesian Networks (BNs) [12], introduced in the 1980s, are Probabilistic Graphical Models that represent uncertain knowledge using Directed Acyclic Graphs (DAGs), where nodes denote random variables and edges represent conditional dependencies. The power of BNs is based on Bayes' Theorem [13] enabling a compact representation of a joint probability distribution over a set of random variables $\mathcal{X} = \{X_1, \dots, X_n\}$. The structure of the DAG encodes conditional independence assumptions, allowing for a computationally tractable factorization of the joint distribution:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{Pa}(X_i)) \quad (1)$$

where $\text{Pa}(X_i)$ denotes the set of parents of variable X_i in the DAG. This factorization significantly reduces the number of parameters required to define the full joint distribution, making even complex probabilistic models feasible.

B. Bayesian Networks in Medical Applications

BNs have been effectively used for clinical triage by modeling relationships between patient symptoms and outcomes. For instance, BN-based systems have been developed for triaging non-traumatic abdominal pain [14] and detecting pediatric asthma with high accuracy [15]. While other studies also explore BNs for triage support, they often highlight the complexity of this approach [16]–[18].

Concurrently, recent advances in computer vision and machine learning have enabled remote, non-contact prediction of vital signs, such as heart rate from eye movements [19], SpO2 levels [20], and blood pressure via rPPG [21]. A recent systematic review and meta-analysis confirms the potential of this approach, showing that data-driven AI models generally outperform conventional triage tools in predicting trauma outcomes [22]–[25]. However, a significant challenge remains to translate the success of these individual models into

a robust integrated system capable of assessing each casualty case as a whole. Relying solely on the isolated predictions of these diverse, often "black-box" neural networks can lead to inaccuracies, particularly in complex environments.

C. Robotics in Search and Rescue

Depending on the context, robotic systems for triage can be broadly divided into remote and contact solutions. Contact approaches, conventionally used in hospitals, such as wearable monitoring devices [26], have a clear limitation: the injured individual must be equipped with the device beforehand, which is often unrealistic in emergency and mass casualty scenarios. This limitation motivates the use of remote solutions. Some approaches [27], [28] use Unmanned Aerial Vehicles (UAVs) to assess casualties from the air, enabling rapid coverage of large or hard-to-reach scenes. However, short flight times, payload limits, and difficulty maintaining a steady viewpoint in wind, precipitation, or clutter make stable data collection challenging. This is especially problematic for video methods that depend on subtle cues, e.g. detecting respiratory distress via abnormal chest motion, because camera or body movement can mask the thoracic signal; such algorithms typically require stationary, stabilized observations to perform reliably.

On the other hand, Unmanned Ground Vehicles (UGVs), allow closer interaction with victims and the use of more sophisticated sensors. However, UGVs face challenges related to terrain accessibility, navigation in debris-filled areas, and the time required to physically approach each casualty.

These limitations highlight that while both UAV and UGV based approaches expand the possibilities of remote triage, neither solution alone fully addresses the complexities of real-world mass casualty incidents. It is important to emphasize that this field is still in its infancy. Current solutions typically focus on either UAVs or UGVs, each addressing only part of the problem. Equally important is the challenge of multimodal data fusion: learning how to aggregate heterogeneous information streams from visual, thermal, acoustic, and physiological sensors into a coherent and reliable triage decision framework. Addressing these challenges will be key to advancing robotic triage systems from experimental prototypes to effective tools useful in real-world incidents.

III. METHODS

A. Environment

Our study was conducted in field tests that simulate MCIs with 20 casualties, constructed for the DTC. Experiments were conducted in accordance with the Institutional Review Board. Autonomous ground and aerial robots used only non-contact, stand-off sensors to locate and assess casualties, estimating vital physiological parameters and visible life-threatening indicators to support early triage. The environment featured unstructured terrain, constrained passages, occlusions, and sensory degradation (fog, dust, background noise); some casualties were partially hidden under debris or inside vehicles, requiring active search. Casualties

TABLE I
SCORING CRITERIA FOR EMERGENCY SCENARIOS.

Field	Values	Scoring Criteria
Severe Hemorrhage	Present / Absent	4 if in GW 2 if match GT 0 otherwise
Respiratory Distress	Present / Absent	4 if in GW 2 if match GT 0 otherwise
Head Trauma	Wound / Normal	2 if all match GT
Torso Trauma		1 if at least 2 match GT
Lower Ext. Trauma	Normal / Wound /	0 otherwise
Upper Ext. Trauma	Amputation	
Ocular Alertness	Open / Closed / NT	2 if all match GT
Verbal Alertness	Normal / Absent /	1 if at least 2 match GT
Motor Alertness	Abnormal / NT	0 otherwise

were human volunteers and TOMManikin [29] high-fidelity manikins, distributed with varied injuries and postures. Our system fused outputs from multiple onboard perception algorithms within a probabilistic reasoning framework to infer physiological state and injury severity from incomplete observations.

System performance was evaluated using a structured scoring framework that awarded point scores for correctly identifying physiological signs and conditions under time constraints. Each casualty presented a set of target findings, and points were awarded when the system accurately reported those findings. In Table I we present a list of physiological measurements together with their corresponding score rubric. Additional weight was given for completing the assessment rapidly, within a “golden time window” (GW), a period in which interventions are likely to prevent death, representing the urgency of real-world triage. The primary objective was to maximize the cumulative score for all casualties during the execution time budget of the scenario.

B. System Architecture and Integration

To address the multifaceted issues involved in the assessment of casualties, our solution leverages a probabilistic modeling framework grounded in domain expertise. The underlying BN model was constructed and processed using GeNIe Modeler [30]. Its seamless integration into the Robot Operating System 2 (ROS 2) [31] stack was achieved through a custom interface developed with SMILE [32], which facilitates real-time inference and data exchange. This implementation follows the principles outlined in [33] for practical application within a modern robotic middleware ecosystem.

The perception system was deployed on a robotic platform equipped with a sensor suite comprising a high-resolution RGB camera, a microphone, a radar, a lidar and a thermal camera enabling the collection of visual and audio data from a stand-off distance. The system is built on the ROS 2

framework, where individual physiological estimators operate as independent nodes. Each node, corresponding to an input variable in the BN model shown in Figure 2, is an independent AI model specialized in assessing a specific physiological sign from the robot sensor feed.

Each estimator node publishes time-stamped predictions (e.g., severe hemorrhage, ocular alertness) to the central BN service. When the robot signals scan completion or at a fixed cadence, the service triggers BN inference to impute unobserved variables and produces a triage posterior that is returned to the scoring module. The explicitly modular sensor/model-agnostic architecture integrates black-box estimator modules, exposing only their predicted variables, without any internal details. Components can therefore be swapped or added without modifying the reasoning layer. This modularity ensures that the system remains adaptable to evolving technologies in emergency response.

The BN treats incoming predictions as uncertain evidence and fuses heterogeneous sources with unknown or variable reliability. Upon receiving inputs, the BN performs the inference process detailed in Equation 1 to estimate unobserved physiological variables and, through this probabilistic treatment, generates a consistent posterior over the patient state. This approach reduces the impact of errors in any single module, allows reasoning over incomplete or noisy inputs, and supports on-the-fly updates as new data become available, making it suitable for real-time triage.

The parameterization of the Bayesian Network, specifically the definition of the Conditional Probability Tables (CPTs) for each node, was carried out through a knowledge-driven engineering process. Due to the absence of publicly available standardized datasets for MCI scenarios, a data-driven approach was not feasible. Instead, the CPTs were populated based on qualitative rules and heuristics elicited through consultations with medical experts, as detailed in Section III-D. This process involved translating qualitative statements (e.g., “highly likely,” “possible,” “unlikely”) into quantitative probability values.

To ensure consistency and facilitate replicability, we adopted the following convention:

- **Strong causal relationships** (e.g., a lower extremity amputation almost always causes severe hemorrhage) were modeled using probabilities in the range of 0.8 – 0.95.
- **Moderate associations** (e.g., head trauma may be associated with abnormal verbal alertness) were assigned values in the 0.4 – 0.6 range.
- **Weak dependencies** were represented by low probabilities, close to the prior probabilities of the baseline.

For instance, the probability for “Closed” for “Ocular Alertness” node in the case when prior of “Head Trauma” was “Wound” was set to 0.7. Although exact values were heuristically assigned, this structured methodology provides transparency and a framework for future model calibration should relevant data become available.

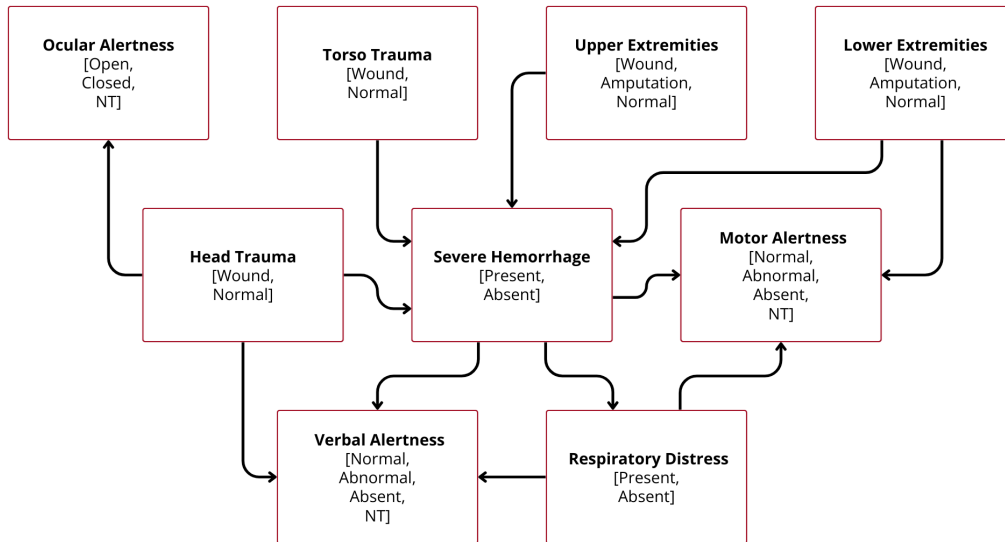


Fig. 2. Bayesian network architecture illustrating the relationships among vital signs.

C. Robotic Platform

The experimental infrastructure is composed of a heterogeneous team of aerial and ground robots operating within a common ROS 2 framework. The robotic fleet consists of three Spot-type quadrupedal ground robots (left UGV in Figure 1), one custom-made aerial platform (top right UAV in Figure 1), and one Indago 4 Scout aerial platform (bottom right UAV in Figure 1). Both UAVs are equipped with GPS, a standard RGB camera, and a thermal camera. Their primary role is to rapidly detect, localize, and track casualties and to detect bleeding, leveraging their ability to cover large areas quickly and provide aerial situational awareness.

The Spot robots are equipped with a diverse set of sensors, including a LiDAR, GPS, multispectral camera, radar, microphone, depth camera, and thermal camera. Their task is twofold: first, to autonomously locate victims on the ground and second, to perform a detailed estimation of vital signs as defined in Table I. This multimodal sensing capability enables the collection of complementary information in detail that is not accessible from aerial platforms.

To ensure consistent and reliable assessment, the predictions generated by each platform are processed through matching algorithms that aggregate observations corresponding to the same casualty. These fused data streams are then forwarded to the BN, which infers any missing information and produces a coherent triage profile. All communication and coordination between platforms are implemented using ROS 2 topics and services, ensuring modularity, scalability, and interoperability across the robotic system. For computations, our robots used NVIDIA 64 GB Jetson AGX Orin.

D. Expert Knowledge Elicitation

The design and parameterization of the BN, whose structure is shown in Figure 2, involved three domain experts,

each with more than a decade of field experience in emergency medicine and rapid response medical interventions.

The methodology consisted of an iterative cycle. Initially, our engineering team conducted a series of individual interviews to identify the most critical observable indicators and their high-level relationships. Based on this qualitative input, an initial BN structure and preliminary Conditional Probability Tables (CPTs) were drafted. These drafts translated qualitative expert statements (e.g., "a lower extremity amputation almost always causes severe hemorrhage") into quantitative probability values.

In the second phase, the drafted model was presented back to the experts in follow-up sessions. Their role was to validate the logic of the model, identify potential gaps, and challenge assumptions. This feedback loop allowed the engineering team to manually refine the CPTs to ensure that the model's inferential behavior was balanced and consistent with the established clinical reasoning. This expert-in-the-loop approach, although not fully automated, ensured that the final model was computationally robust and grounded in real-world medical expertise.

IV. RESULTS

The effectiveness of our framework was evaluated during the DTC Round 1 Systems Competition. We report performance results for two scenarios. The third run is excluded from this analysis due to a hardware-level failure prior to deployment that prevented consistent data acquisition. Despite this, Team Chiron's performance in the remaining missions was sufficient to secure 4th place out of 11 teams overall. Our findings demonstrate a significant improvement in the robot's triage capabilities when equipped with our Bayesian reasoning engine. The comparative analysis between BN-based cognitive architecture ("Robot + BN"), against a baseline system that relies solely on the direct outputs of its

TABLE II
COMPARISON OF THE SYSTEM PERFORMANCE WITH AND WITHOUT BN

Casualty Id	Run A Scores		Run B Scores	
	Robot	Robot + BN	Robot	Robot + BN
1	9	9	0	0
2	0	3	0	3
3	0	3	2	7
4	0	3	2	7
5	0	3	0	4
6	0	5	1	3
7	4	5	8	11
8	0	7	1	3
9	8	9	2	7
10	4	7	-	-
11	0	7	-	-
Total	25/132	61/132	16/108	45/108

perception pipeline ("Robot") on 20 distinct casualty cases divided into two separate operationally realistic scenarios termed Run A and Run B are detailed in Table II. Our integrated system more than doubled the baseline score in Run A (from 25 to 61) and nearly tripled it in Run B (from 16 to 45). This highlights the architecture's ability to overcome the limitations of real-world robotic perception, where sensor failures, environmental occlusions, and ambiguous visual cues are common.

Table III further details this improvement across three key metrics. System's **Reliability**, which measures the robots' ability to provide an assessment even with partial data, increased from 0.31 to 0.95. This indicates that the baseline failed to make an assessment in 69% of cases, whereas our integrated system provided an output in 95% of them. Achieving 100% reliability was not achievable due to the inability of the system to locate one of the casualties in Run B. Furthermore, the quality of these assessments improved, with **Accuracy** (correct assignments among those attempted) rising from 46% to 56%.

TABLE III
COMPARISON OF CORRECT ASSIGNMENTS AND ASSIGNMENT ATTEMPTS WITH AND WITHOUT BN ACROSS ALL CASUALTIES.

Vital	Correct Assignments		Assignment Attempts	
	Robot	Robot + BN	Robot	Robot + BN
Severe Hemorrhage	6	12	12	19
Respiratory Distress	5	16	6	19
Head Trauma	0	15	0	19
Torso Trauma	0	11	0	19
Lower Ext. Trauma	8	11	14	19
Upper Ext. Trauma	3	8	14	19
Ocular Alertness	0	8	1	19
Motor Alertness	0	8	0	19
Verbal Alertness	3	7	8	19
Total	25	96	55	171
Reliability	-	-	0.31	0.95
Performance	14%	53%	-	-
Accuracy	46%	56%	-	-

Most importantly, the overall **Performance**, which measures the rate of correct assignments across all possible vitals, increased nearly four-fold from 14% to 53%. This was largely attributable to the BN's ability to infer missing information and resolve conflicting perceptual cues, resulting in a 3.84-fold increase in the total number of correctly assigned vitals. This underscores the critical role of the cognitive architecture in transforming the fragmented, unreliable outputs of a standard robotic perception pipeline into a coherent and actionable assessment, a crucial step for deploying autonomous systems in high-stakes scenarios.

Figure 3 presents a comparative analysis of performance accuracy, distinguishing between configurations with and without the integration of the BN, relative to a random classifier baseline that uniformly selects from the set of valid labels. The inclusion of the BN demonstrably enhanced the system's performance, enabling it to surpass the random classifier across all evaluated vital signs.

A. Real-time Performance

One design objective was to ensure our solution could operate in real time on field-deployable hardware. Our implementation achieves inference times of less than **1 millisecond** per update and maintains a memory footprint below **100 megabytes**. This performance enables deployment on resource-constrained platforms, such as the NVIDIA Jetson series or even a Raspberry Pi, without compromising the robot's responsiveness or reliability in real-world scenarios.

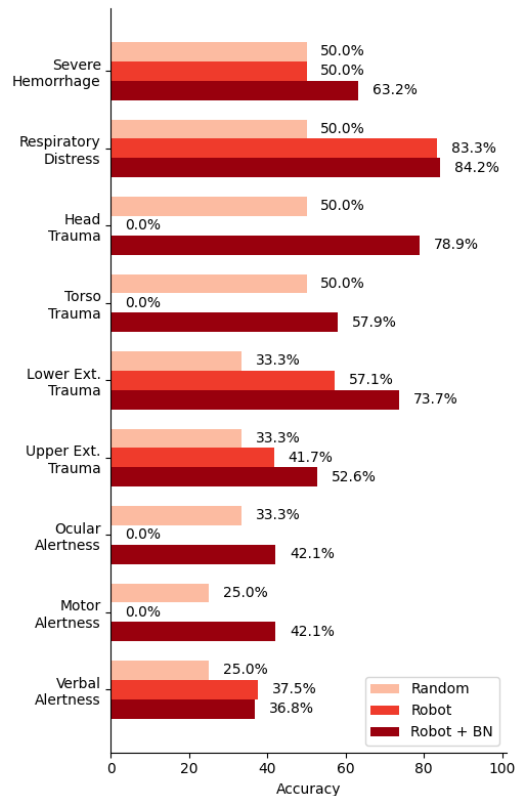


Fig. 3. Comparison of random classifier, robot and robot + BN

V. DISCUSSION

Our results demonstrate that integrating a Bayesian reasoning engine with advanced vision-based sensing can produce a substantial improvement in the casualty assessment capabilities of an autonomous robotic system operating in the field. This highlights the profound impact of explicitly modeling uncertainty and intermodal dependencies, a capability that standard robotic perception pipelines lack. The system's performance improves as it gathers more evidence, emulating the iterative reasoning of human responders at machine speeds. This shift from static, one-off assessments to continuous probabilistic reasoning is a key step towards building truly effective and useful autonomous triage robots.

This research underscores a critical limitation of many current robotic systems: reliance on isolated perception modules, which often function as "black boxes" in isolation from domain knowledge. In the high-stakes, chaotic environment of an MCI, this can lead to inconsistent or medically implausible predictions. For example, a robot's perception pipeline might correctly identify a limb amputation while another module, seeing minor movements, incorrectly reports normal motor alertness. Our cognitive architecture, centered on the Bayesian Network, resolves such contradictions by enforcing clinical constraints. This offers a key advantage in system-level robustness and is paramount for any robot deployed in a safety-critical role.

A fundamental requirement for any field-deployable robot is fault tolerance. Our architecture enables graceful degradation when perceptual data is incomplete. A failure of a single input module, a common occurrence due to sensor occlusions, communication loss, or algorithmic errors, does not cause a catastrophic failure of the overall case assessment task. Instead, the BN marginalizes over the missing variable, providing the best possible assessment based on the available data from all other functioning sensors. This operational reliability is essential for autonomous systems operating outside controlled lab environments.

Furthermore, the probabilistic framework is inherently extensible, making the system's cognitive architecture future-proof. Although the current implementation treats categorical inputs as direct evidence, the BN structure readily accommodates future enhancements. As new sensors or improved perception algorithms are added to the system, their outputs can be integrated with minimal changes to the core reasoning engine. It could also be extended to incorporate model confidence scores or to implement mechanisms for rejecting spurious sensor readings, further enhancing robustness. This forward-looking flexibility was a key factor in our architectural design.

Although our initial results are encouraging, a more comprehensive evaluation is in order. Future work should involve direct comparisons with existing deterministic and probabilistic approaches, including rule-based expert systems, hybrid neuro-symbolic models, and hybrid deep learning pipelines.

VI. CONCLUSION AND FUTURE WORK

In this work, we introduced a cognitive architecture to address the critical challenge of robust decision making for autonomous robots deployed in mass casualty incidents. By centering our system on an expert-guided Bayesian Network, we enabled it to fuse fragmented and uncertain data from its perception pipeline into a coherent multifactor and medically plausible triage assessment. The field evaluations conducted during the DTC Round 1 demonstrated that this architecture can substantially improve reliability and overall performance, transforming the presented system from a simple data collector into an effective decision-making agent. This work represents a key step towards deploying autonomous systems in safety-critical real-world scenarios.

Based on our findings, we propose three primary directions for future research to build upon the foundation presented.

A. Human-Robot Teaming and Ethical Deployment

In any realistic deployment, the robots will operate as part of a team with human responders. Future research must focus on the principles of effective human-robot teaming. This includes designing intuitive interfaces for communicating the system-generated assessments and corresponding uncertainties to a human operator, enabling shared situational awareness and collaborative decision making. As with any autonomous system operating in high-stakes contexts, ethical considerations are paramount. Future development must incorporate robust safeguards and human-in-the-loop oversight mechanisms to ensure that autonomous triage systems are deployed responsibly and effectively as tools to aid, not replace, human experts.

B. Robustness Testing

Our experiments have demonstrated a significant improvement in key metrics, confirming the effectiveness of the presented approach. However, a fundamental conclusion drawn from our analysis is the critical importance of robustness testing. This is especially critical given the model's multi-modal architecture, which relies on inputs generated by other models. Such structures are susceptible to cascading error propagation, where even minor inaccuracies in one of the input modules could lead to significant errors in the final output, even though we have not observed such effects in our experiments.

C. Data Quality and Availability.

A fundamental challenge in this work was the complete absence of empirical data that could be used for machine learning. In our application domain, there are no relevant publicly available standardized datasets. In addition, the specific nature of the triage task imposes extremely high data quality requirements. Any error in the classification process can lead to severe consequences and untoward patient outcomes; therefore, any training dataset would need to be free from annotation errors. We are pushing to collect relevant data during our field experiments, but we would like to encourage the research community to join us in these efforts.

ACKNOWLEDGMENT

This work has been partially supported by the Defense Advance Research Projects Agency (award HR00112420329) and National Science Foundation (awards 2427948 and 2406231). The views, opinions, and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. We also thank the DARPA Triage Challenge organizers and the technical observers for their logistical support during the field evaluations

REFERENCES

- [1] R. Gunaratna and C. Haynal, "Current and emerging threats of homegrown terrorism: The case of the Boston bombings," *Perspectives on Terrorism*, vol. 7, no. 3, pp. 44–63, 2013.
- [2] A. Tahernejad, A. Sahebi, A. S. S. Abadi, and M. Safari, "Application of artificial intelligence in triage in emergencies and disasters: a systematic review," *BMC Public Health*, vol. 24, no. 1, p. 3203, 2024.
- [3] G. Super, S. Groth, R. Hook *et al.*, "START: simple triage and rapid treatment plan," *Newport Beach, CA: Hoag Memorial Presbyterian Hospital*, vol. 199, 1994.
- [4] B. J. Gabbe, W. Veitch, A. Mather, K. Curtis, A. J. Holland, D. Gomez, I. Civil, A. Nathens, M. Fitzgerald, K. Martin *et al.*, "Review of the requirements for effective mass casualty preparedness for trauma systems. A disaster waiting to happen?" *British journal of anaesthesia*, vol. 128, no. 2, pp. e158–e167, 2022.
- [5] K. Hashimoto, T. Matsuzawa, T. Teramachi, K. Uryu, X. Sun, S. Hamamoto, A. Koizumi, and A. Takanishi, "A four-limbed disaster-response robot having high mobility capabilities in extreme environments," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 5398–5405.
- [6] S. Agarwal, R. R. Murphy, and J. A. Adams, "Characteristics of indoor disaster environments for small uass," in *2014 IEEE International Symposium on Safety, Security, and Rescue Robotics (2014)*. IEEE, 2014, pp. 1–6.
- [7] N. Ádám, D. Val"Ko, and E. Chovancová, "Using modern hardware and software solutions for mass casualty incident management," in *2023 IEEE 27th International Conference on Intelligent Engineering Systems (INES)*. IEEE, 2023, pp. 000 183–000 188.
- [8] K. Tiwari, X. Xiao, A. Malik, and N. Y. Chong, "A unified framework for operational range estimation of mobile robots operating on a single discharge to avoid complete immobilization," *Mechatronics*, vol. 57, pp. 173–187, 2019.
- [9] D. Patel, P. Pham, K. Tiwari, and A. Bera, "Dream: Decentralized reinforcement learning for exploration and efficient energy management in multi-robot systems," *arXiv preprint arXiv:2309.17433*, 2023.
- [10] S. Park, Y. Oh, and D. Hong, "Disaster response and recovery from the perspective of robotics," *International Journal of Precision Engineering and Manufacturing*, vol. 18, no. 10, pp. 1475–1482, 2017.
- [11] J. J. Wadden, "Defining the undefinable: the black box problem in healthcare artificial intelligence," *Journal of Medical Ethics*, vol. 48, no. 10, pp. 764–768, 2022.
- [12] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Elsevier, 2014.
- [13] T. Bayes, "LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S," *Philosophical transactions of the Royal Society of London*, no. 53, pp. 370–418, 1763.
- [14] S. Sadeghi, A. Barzi, N. Sadeghi, and B. King, "A Bayesian model for triage decision support," *International journal of medical informatics*, vol. 75, no. 5, pp. 403–411, 2006.
- [15] D. L. Sanders and D. Aronsky, "Prospective evaluation of a Bayesian network for detecting asthma exacerbations in a pediatric emergency department," in *AMIA Annual Symposium Proceedings*, vol. 2006, 2006, p. 1085.
- [16] M. M. Abad-Grau, J. Ierache, C. Cervino, and P. Sebastiani, "Evolution and challenges in the design of computational systems for triage assistance," *Journal of biomedical informatics*, vol. 41, no. 3, pp. 432–441, 2008.
- [17] M. Fareh, I. Riali, H. Kherbache, and M. Guemouz, "Probabilistic reasoning for diagnosis prediction of Coronavirus disease based on probabilistic ontology," *Computer Science and Information Systems*, vol. 20, no. 3, pp. 1109–1132, 2023.
- [18] R. T. Olszewski, "Bayesian Classification of Triage Diagnoses for the Early Detection of Epidemics." in *FLAIRS*, 2003, pp. 412–416.
- [19] K. Zheng, K. Ci, H. Li, L. Shao, G. Sun, J. Liu, and J. Cui, "Heart rate prediction from facial video with masks using eye location and corrected by convolutional neural networks," *Biomedical Signal Processing and Control*, vol. 75, p. 103609, 2022.
- [20] A. Al-Naji, G. A. Khalid, J. F. Mahdi, and J. Chahl, "Non-contact SpO₂ prediction system based on a digital camera," *Applied Sciences*, vol. 11, no. 9, p. 4255, 2021.
- [21] F. Schruppf, P. Frenzel, C. Aust, G. Osterhoff, and M. Fuchs, "Assessment of non-invasive blood pressure prediction from ppg and rppg signals using deep learning," *Sensors*, vol. 21, no. 18, p. 6022, 2021.
- [22] M. Yuba and K. Iwasaki, "Systematic analysis of the test design and performance of AI/ML-based medical devices approved for triage/detection/diagnosis in the USA and Japan," *Scientific reports*, vol. 12, no. 1, p. 16874, 2022.
- [23] O. Adebayo, Z. A. Bhuiyan, and Z. Ahmed, "Exploring the effectiveness of artificial intelligence, machine learning and deep learning in trauma triage: A systematic review and meta-analysis," *Digital health*, vol. 9, p. 20552076231205736, 2023.
- [24] A. Vântu, A. Vasilescu, and A. Băicoianu, "Medical emergency department triage data processing using a machine-learning solution," *Heliyon*, vol. 9, no. 8, 2023.
- [25] A. Defilippo, G. Bertucci, C. Zurzolo, P. Veltri, and P. H. Guzzi, "On the computational approaches for supporting triage systems," *Interdisciplinary Medicine*, vol. 1, no. 3, p. e20230015, 2023.
- [26] C. Polley, T. Jayarathna, U. Gunawardana, G. Naik, T. Hamilton, E. Andreozzi, P. Bifulco, D. Esposito, J. Centracchio, and G. Gargiulo, "Wearable bluetooth triage healthcare monitoring system," *Sensors*, vol. 21, no. 22, p. 7586, 2021.
- [27] C. Álvarez-García, S. Cámara-Anguita, J. M. López-Hens, N. Granero-Moya, M. D. López-Franco, I. María-Comino-Sanz, S. Sanz-Martos, and P. L. Pancorbo-Hidalgo, "Development of the aerial remote triage system using drones in mass casualty scenarios: a survey of international experts," *PLoS One*, vol. 16, no. 5, p. e0242947, 2021.
- [28] J. Lu, X. Wang, L. Chen, X. Sun, R. Li, W. Zhong, Y. Fu, L. Yang, W. Liu, and W. Han, "Unmanned aerial vehicle based intelligent triage system in mass-casualty incidents using 5G and artificial intelligence," *World journal of emergency medicine*, vol. 14, no. 4, p. 273, 2023.
- [29] North American Rescue Simulation and Techline Technologies, "TOMManikin@ trauma manikin simulator," 2025, accessed 2025-09-14. [Online]. Available: <https://www.narescue.com/tommanikin-trauma-manikin-simulator>
- [30] BayesFusion, LLC, "GeNle Modeler," 2022, [Computer software]. [Online]. Available: <https://www.bayesfusion.com/genie/>
- [31] Y. Maruyama, S. Kato, and T. Azumi, "Exploring the performance of ROS2," in *Proceedings of the 13th international conference on embedded software*, 2016, pp. 1–10.
- [32] BayesFusion, LLC, "SMILE Engine," 2022, [Computer software]. [Online]. Available: <https://www.bayesfusion.com/smile/>
- [33] M. J. Druzdzel, "SMILE: Structural Modeling, Inference, and Learning Engine and GeNle: a development environment for graphical decision-theoretic models," in *Aaai/Iaai*, 1999, pp. 902–903.