KHAIT: K-9 Handler Artificial Intelligence Teaming for Collaborative Sensemaking

Matthew Wilchek[1]*, Linhan Wang[1], Feras A. Batarseh[2], and Kurt Luther[1]

[1]Department of Computer Science [2]Department of Biological Systems Engineering
Virginia Tech
(Will attend the workshop)*

Keywords:
Sensemaking, Social Computing, Augmented Reality, Human-in-the-loop, Object Detection, Shared Perception, Animal-Computer Interaction

Abstract:
Following natural and manmade disasters, urban search and rescue (USAR) involving canines (i.e., K-9) face time delays and communication hurdles. This paper proposes a method to improve USAR sensemaking by incorporating object detection Artificial Intelligence (AI) and Augmented Reality (AR), aiming to close the communication gap between USAR animals and their handlers. SAR dogs are equipped with an AI-powered camera and edge computing hardware, while their handlers wear HoloLens-2 AR headsets. This setup allows for advanced object detection from the dogs' perspective, with captured data transmitted to handlers' headsets as hologram indicators to enhance survivor localization. This multidisciplinary strategy, merging human-in-the-loop (HITL), canine, and AI capabilities, aims to transform SAR missions, ensuring faster and safer rescues. At the CHI sensemaking workshop, we hope to solicit feedback regarding our proposed design and evaluation that will have broad relevance to attendees interested in human-AI and animal-computer interaction, AR, and/or disaster response.

Extended Abstract

Introduction: USAR Sensemaking with Canines

In the wake of manmade and natural disasters (e.g., earthquakes, tornadoes, wars), sensemaking, a decision-making process that involves creating plausible explanations for human actions, becomes vital to recovery missions [1, 2]. The complex effort to search and locate trapped survivors from collapsed structures or debris is a race against the clock. Rescuers who conduct this life-saving activity specialize in urban search and rescue (USAR). USAR teams are composed of highly trained professionals from various fields, including firefighters, engineers, medical personnel, and canine search handlers. Canine search teams are trained to work and detect in any environment and are not limited by noise, equipment, or distractions. Due to their heightened sense of smell, dogs can detect live human scent and can specialize in detecting survivors or the deceased. According to the US Federal Emergency Management Agency (FEMA), there are 284 canine search teams dedicated to searching for post-disaster survivors, and 90 teams that search for human remains, as of March 2020 [3].

When a canine detects a survivor, it initiates traditional methods of communication to notify the handler, including barking at the site or returning to the handler and leading them to the survivor's location [4]. These delays could result in critical minutes lost from applying triage to a survivor and saving their life. A handler must quickly make sense of the situation and potentially overcome navigational difficulties within crumbled buildings, communication limitations due to obstructed lines of sight, fire, smoke, low visibility, and psychological stress, which all can affect the decision-making and efficiency of swiftly locating the victim.

Augmented Reality (AR) and edge computational hardware have emerged as key enablers for such extreme sensemaking situations [5, 6]. AR devices use onboard cameras to capture the environment, which aids in object detection and localization in three-dimensional space, thereby significantly enhancing situational awareness. In recent years, the exploration of robotics in USAR operations has necessitated the integration of various sensors, edge computing capabilities, and computer vision algorithms [7, 8]. By drawing upon these insights, we see potential in integrating these technological advancements with AR, specifically employing object detection AI trained to identify hard-to-find individuals, such as those partially occluded. This approach not only fosters human-AI and canine-AI collaborations, but also paves the way for novel AI teaming, leveraging the unique strengths of both humans and canines in a collaborative sensemaking framework for disaster response scenarios.

Related Work

Human-Computer Interaction (HCI) and Animal-Computer Interaction (ACI) represent two complementary schools of thought in the design of interactive systems [9, 10]. Integration of AR technology for SAR operations is highly dependent on its compatibility with the specialized equipment used by professionals, as illustrated by the development of ruggedized HoloLens adaptations such as the C-THRU helmet, IVAS, and MARS for enhanced field applicability [11, 12, 13]. These projects underscore the potential of AR to significantly improve sensemaking [14, 15]. For instance, Smith et al.'s 2022 software for HoloLens aids in bridge inspections through annotations and spatial markers [16], while Guan et al. introduced an object detection app that overlays 3D bounding boxes on detections [17], furthering the use of AR headsets' depth data in object detection [18, 19, 20].

Related work in our context for animal-centered design is less extensive. One example, from Alcaidinho et al. [21] showcases wearable technology for canine units, featuring sensors and GPS on a harness to relay critical events, such as explosive detection, to officers through a ruggedized cellphone. Pai et al. [22] introduces another smart guide dog harness utilizing AI edge computing for assisting visually impaired individuals, integrating image recognition and navigation technologies. Conversely, Kasnesis et al. [23] leverages a different wearable-based approach for SAR canines, utilizing deep Convolutional Neural Networks (CNN) to detect activities and barking in real-time by analyzing data from inertial sensors and microphones. The system shows promising results in improving communication between SAR canines and their handlers by providing immediate feedback on the canine's status and findings. While these systems are notable for their contributions in wearable tech and activity detection, they focus on isolated aspects like communication or navigation, lacking comprehensive integration. Recognizing SAR's need for intuitive visual communication and hands-free navigation, we aimed to create a system merging visual cues, AI detection, and AR to enhance canine-handler coordination and situational awareness.

System Description: KHAIT

User Scenario: As mentioned above, traditional search methods in partially collapsed structures involve canines locating potential survivors and signaling their handlers through barking. The handler, relying solely on auditory cues, must then navigate the debris to reach the dog’s location. With our system, which we call KHAIT (K-9 Handler Artificial Intelligence Teaming), when the canine discovers a survivor, the canine’s location is instantly transmitted to the handler’s wearable AR device. This information enables the handler to navigate to the canine and survivor’s location swiftly and efficiently, guided by a hologram marker visible through the AR headset. This marker, anchored to the precise location of the dog or survivor, remains visible even through physical obstructions. Finally, the handler reaches the marker and triages the survivor, and/or request specific types of support as needed. Figure 1 depicts a rough design where the animal handler watches in AR while being detected by the AI-edge harness.

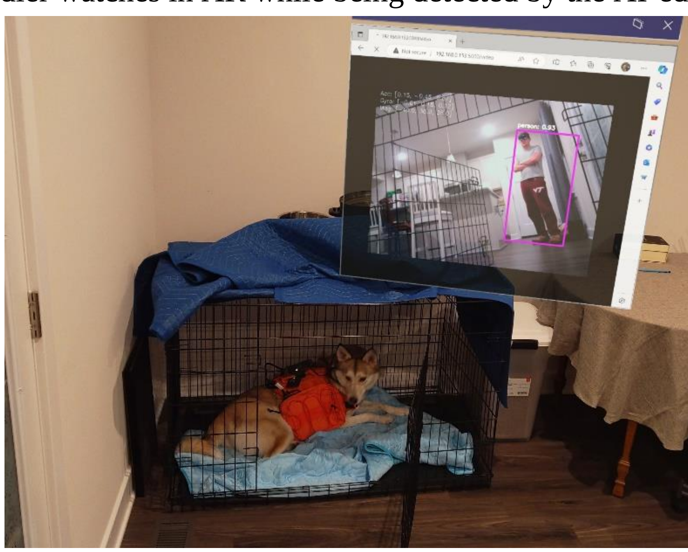

Figure 1: Photograph of the initial KHAIT prototype.

Implementation Details: KHAIT differentiates itself from earlier sensemaking technologies by incorporating visual cues delivered via a wearable AR headset. The wearable harness we developed, illustrated in Figure 2, is equipped with an NVIDIA Jetson Orin Nano (8GB) for computational tasks, an Arduino Nano 33 Sense BLE, a CMOS 4K Autofocus Camera for high-resolution imaging, and a 27000mAh portable power bank to ensure extended operational duration.

The Jetson Orin Nano processes an advanced embedded object detection algorithm specifically trained to identify people. This critical data, along with Inertial Measurement Unit (IMU) readings, is relayed directly to an AR headset worn by the USAR canine handler through socket programming over a secure local network. To initially calibrate the system, initial IMU data from the AR headset is shared with the Arduino Nano, establishing a common understanding of a single point in a mutually understood physical space. This process allows for the measurement of relative distance from this waypoint as users move away from it, facilitating collaborative, accurate, and large-scale collaboration in AR without the need for QR codes or referencing imagery. Figure 3 illustrates an example of what the detection may look like to the handler.

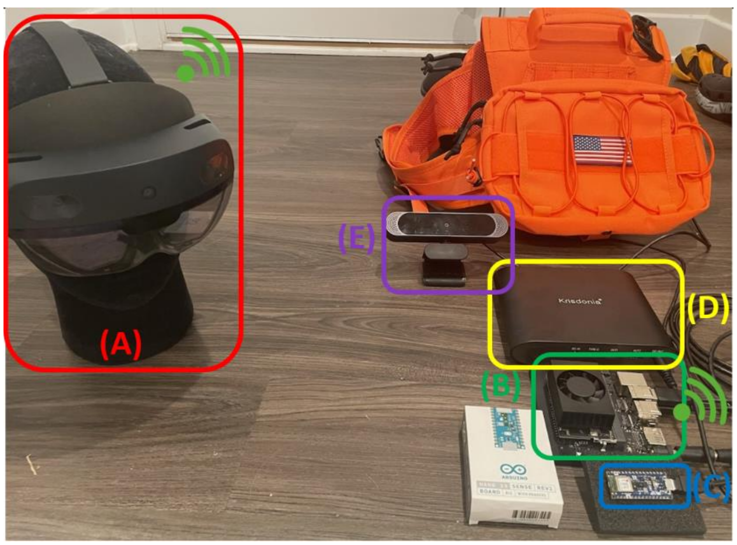

Figure 2: Photograph of the KAIT prototype. (A) HoloLens-2 for animal handler. (B) AI edge computing module. (C) Arduino Nano 33 Sense board for IMU data. (D) Power bank. (E) CMOS 4k Autofocus camera.

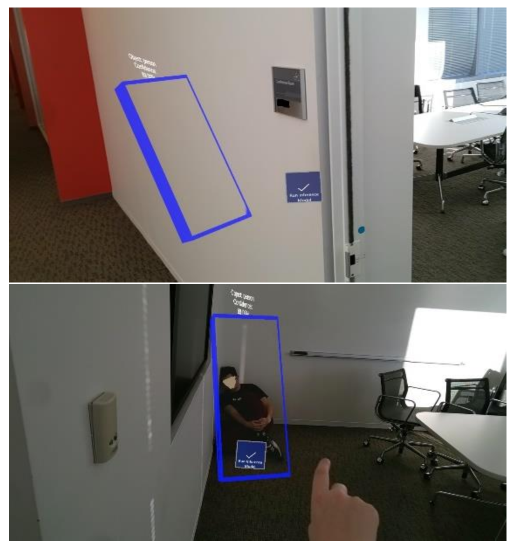

Figure 3: Examples of object detection results sent via canine and seen by handler wearing HoloLens-2.

Proposed Evaluation

Following preliminary testing, our next step is to conduct user studies with SAR professionals, especially those experienced in canine search, to evaluate KHAIT. These studies may simulate search and rescue missions, using a canine to locate a “survivor” actor, aiming to assess the system's efficiency by timing how quickly a rescuer can reach the identified location. Our evaluation approach includes both quantitative and qualitative methods. Quantitatively, we will measure the time reduction in locating survivors versus traditional methods reliant on canine sounds. Qualitatively, we aim to study sensemaking cognition and behavior and gather feedback on the usability and feasibility of equipping SAR canines and handlers with our technology. At the CHI workshop, we hope that attendees’ areas of interest, expertise, and feedback would be most helpful in advancing our understanding of how KHAIT can not only improve operational efficiency in USAR scenarios but also contribute to the broader field of collaborative sensemaking.

References

[1] K. E. Weick, Sensemaking in Organizations, Sage Publications, 1995.
[2] S. M. Fiore, T. J. Wiltshire, R. A. Lashlee, E. Salas, Distributed sensemaking: A case study of military analysis, *Journal of Organizational Behavior 31* (2010) 291–307. doi:10.1002/job.619.
[3] Federal Emergency Management Agency (FEMA), Canines' role in urban search & rescue, https://www.fema.gov/emergency-managers/national-preparedness/frameworks/urban2021. Accessed: 2024-02-24.
[4] Tactical Police K9 Training, How do search and rescue dogs locate victims of disasters?, 2021. URL: https://tacticalpolicek9training.com/how-do-search-and-rescue-dogs-locate-victims-of-disasters/
[5] M. Billinghurst, A. Clark, G. Lee, A survey of augmented reality, *Foundations and Trends® in Human–Computer Interaction 8* (2015) 73–272. doi:10.1561/1100000049.
[6] V. Krauß, A. Boden, L. Oppermann, R. Reiners, Current practices, challenges, and design implications for collaborative ar/vr application development, in: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, 2021, pp. 1–15. doi:10.1145/3411764.3445335.
[7] R. Ventura, P. U. Lima, Search and rescue robots: The civil protection teams of the future, in: *2012 Third International Conference on Emerging Security Technologies*, 2012, pp. 12–19. doi:10.1109/EST.2012.40.
[8] J. G. Blitch, R. Maurer, Knobsar: A knowledge based system prototype for robot assisted urban search and rescue, *SIMULATION 66 (1996)* 375–391. URL: https://doi.org/10.1177/003754979606600605.
[9] B. Shneiderman, Human-centered AI, Oxford University Press, 2022.
[10] J. Farrell, C. McCarthy, C. Chua, Adapting hci techniques for the design and evaluation of canine training technologies, in: *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, OzCHI '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 189–193. URL: https://doi-org.ezproxy.lib.vt.edu/10.1145/3292147.3292191. doi:10.1145/3292147.3292191.
[11] Qwake Technologies, Qwake-technologies c-thru, n.d. URL: https://www.qwake.tech/, accessed: 2024-02-24.
[12] Shear, F. (2023, August 1). Army accepts prototypes of the most advanced version of IVAS. Retrieved February 24, 2024, from https://www.army.mil/article/268702/army_accepts_prototypes_of_the_most_advanced_version_of_ivas
[13] Keller, J. (2023, January 27). China's military unveils heads-up display to let soldiers shoot around corners: Meet the MARS, the Chinese military's IVAS clone. Retrieved February 24, 2024, from https://taskandpurpose.com/tech-tactics/china-military-augmented-reality-system-weapons/
[14] A. Dasgupta, M. Manuel, R. S. Mansur, N. Nowak, D. Graˇcanin, Towards real time object recognition for context awareness in mixed reality: A machine learning approach, in: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2020, pp. 262–268. doi:10.1109/VRW50115.2020.00054.

[15] M. Hu, D. Weng, F. Chen, Y. Wang, Object detecting augmented reality system, in: *2020 IEEE 20th International Conference on Communication Technology (ICCT)*, 2020, pp. 1432–1438.doi:10.1109/ICCT50939.2020.9295761.
[16] A. Smith, C. Duff, R. Sarlo, J. L. Gabbard, Wearable augmented reality interface design for bridge inspection, in: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2022, pp. 497–501. doi:10.1109/VRW55335.2022.00111.
[18] A. Ahmadyan, L. Zhang, J. Wei, A. Ablavatski, M. Grundmann, Objectron: A large-scale dataset of object-centric videos in the wild with pose annotations, *CoRR abs/2012.09988 (2020)*. URL: https://arxiv.org/abs/2012.09988.
[19] J. Lahoud, B. Ghanem, 2d-driven 3d object detection in rgb-d images, in: *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 4632–4640. doi:10.1109/ICCV.2017.495.
[20] H. Ye, J. Leng, C. Xiao, L. Wang, H. Fu, Proobjar: Prototyping spatially-aware interactions of smart objects with ar-hmd, in: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, 2023, pp. 1–15. doi:10.1145/3544548.3580750.
[21] J. Alcaidinho, L. Freil, T. Kelly, K. Marland, C. Wu, B. Witten brook, G. Valentin, M. Jackson, Mobile collaboration for human and canine police explosive detection teams, in: *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17*, Association for Computing Machinery, New York, NY, USA, 2017, p. 925–933. URL: https://doi-org.ezproxy.lib.vt.edu/10.1145/2998181.2998271. doi:10.1145/2998181.2998271.
[22] W. Y. Pai, W.-H. Chen, L.-B. Chen, An ai edge computing-based smart guide dog harness for assisting the visually impaired, in: *2023 IEEE 12th Global Conference on Consumer Electronics (GCCE)*, 2023, pp. 1071–1072. doi:10.1109/GCCE59613.2023.10315353.
[23] P. Kasnesis, V. Doulgerakis, D. Uzunidis, D. G. Kogias, S. I. Funcia, M. B. Gonz´alez, C. Giannousis, C. Z. Patrikakis, Deep learning empowered wearable-based behavior recognition for search and rescue dogs, *Sensors 22 (2022)*. URL: https://www.mdpi.com/1424-8220/22/3/993. doi:10.3390/s22030993.