Intelligent Autonomous Robotic Car for Real-Time Disaster Area Analysis and Navigation

E.A. Wanigasekara[1], Y.A.A. Kumarayapa[2]

[1,2]Department of Electronics, Wayamba University of Sri Lanka, Kuliyapitiya, Sri Lanka

DOI: https://dx.doi.org/10.51584/IJRIAS.2025.10120082

Received: 29 December 2025; Accepted: 03 January 2026; Published: 17 January 2026

ABSTRACT

Efficient victim detection and reliable navigation remain major challenges in robotic search and rescue operations within disaster affected regions. This research describes the design and implementation of an AI-driven autonomous robot car capable of making real-time decisions in complex and hazardous environments. The proposed system employs a sensor fusion approach that combines visual human detection using YOLOv5, thermal-based classification through a convolutional neural network, and audio-based human voice detection. These AI modules are supported by additional sensors including ultrasonic sensors, INMP441 microphone, MPU6050 inertial unit, and gas sensors (MQ2 and MQ135), all coordinated using a Raspberry Pi 3B+, ESP32, and ESP32-CAM modules. Precise localization and remote communication are achieved using a NEO-6M GPS receiver and a SIM800L GSM module. A web-based monitoring platform is developed to display real-time sensor readings, survivor locations, and environmental hazard warnings at a base station. The system is validated using a physical prototype designed for low-cost, rapid deployment, and ease of use. Experimental observations indicate that the robot can autonomously navigate, identify potential survivors, and transmit critical information, highlighting its suitability for disaster-response applications.

Keywords: Autonomous Robot, Disaster Response, Sensor Fusion, Real-time

INTRODUCTION

Current search and rescue methodologies are constrained by human physical limitations, safety protocols, and delays inherent in site assessment. Natural disasters such as earthquakes and building collapses frequently result in survivors being trapped in hazardous and inaccessible environments. Traditional search and rescue operations in these scenarios are often time-consuming, resource-intensive, and pose significant risks to first responders who must navigate unstable debris, toxic atmospheres, and unpredictable structural conditions. The urgency of post-disaster rescue efforts is heightened by the exponential decrease in survivors' chances of survival over time, necessitating innovative solutions that accelerate search operations while reducing human risk[1]. Although robotic systems have been investigated to aid rescuers, many existing platforms are costly, require specialized training, or depend on infrastructural support unavailable in resource-limited, disaster-prone developing regions[2]. Furthermore, conventional robotic solutions often rely on remote control or pre-programmed behaviors, which restricts their adaptability and real-time decision-making capabilities. Most are limited using single-sensor modalities, such as visual cameras, which fail to provide sufficient situational awareness in complex environments characterized by poor visibility and hazardous conditions[3]. This paper presents the development of an AI-driven autonomous robotic car designed for real-time decision-making in disaster zones. By integrating multiple sensor modalities, including visual cameras, thermal imaging, audio detection, gas sensing, and GPS positioning with advanced AI models such as YOLOv5, convolutional neural networks, and audio classifiers, the system achieves comprehensive environmental perception and adaptive navigation. The design prioritizes low-cost and modularity, making it accessible for deployment in developing countries with dense and complex urban structures. Through this approach, the proposed system aims to enhance the efficiency and safety of post-disaster search and rescue missions [4].

METHODOLOGY

The proposed AI-driven autonomous robot car was developed on a modular embedded platform combining Raspberry Pi 3B+, ESP32, and ESP32-CAM modules for computation and sensor management. The robot integrates multiple sensing modalities. Visual detection is performed using YOLOv5 for human identification from disaster imagery. Thermal detection is performed using a CNN (convolutional neural network) based classifier for differentiating human heat signatures. Audio detection is done by employing a trained classifier on INMP441 microphone inputs to recognize human voices[5]. The design integrates Raspberry Pi, Arduino Uno, ESP32, GPS module, motor drivers, MQ-2 and MQ-135 gas sensors, ultrasonic sensors, PIR sensor, and supporting components for real-time hazard detection and control as shown in Figure 1.

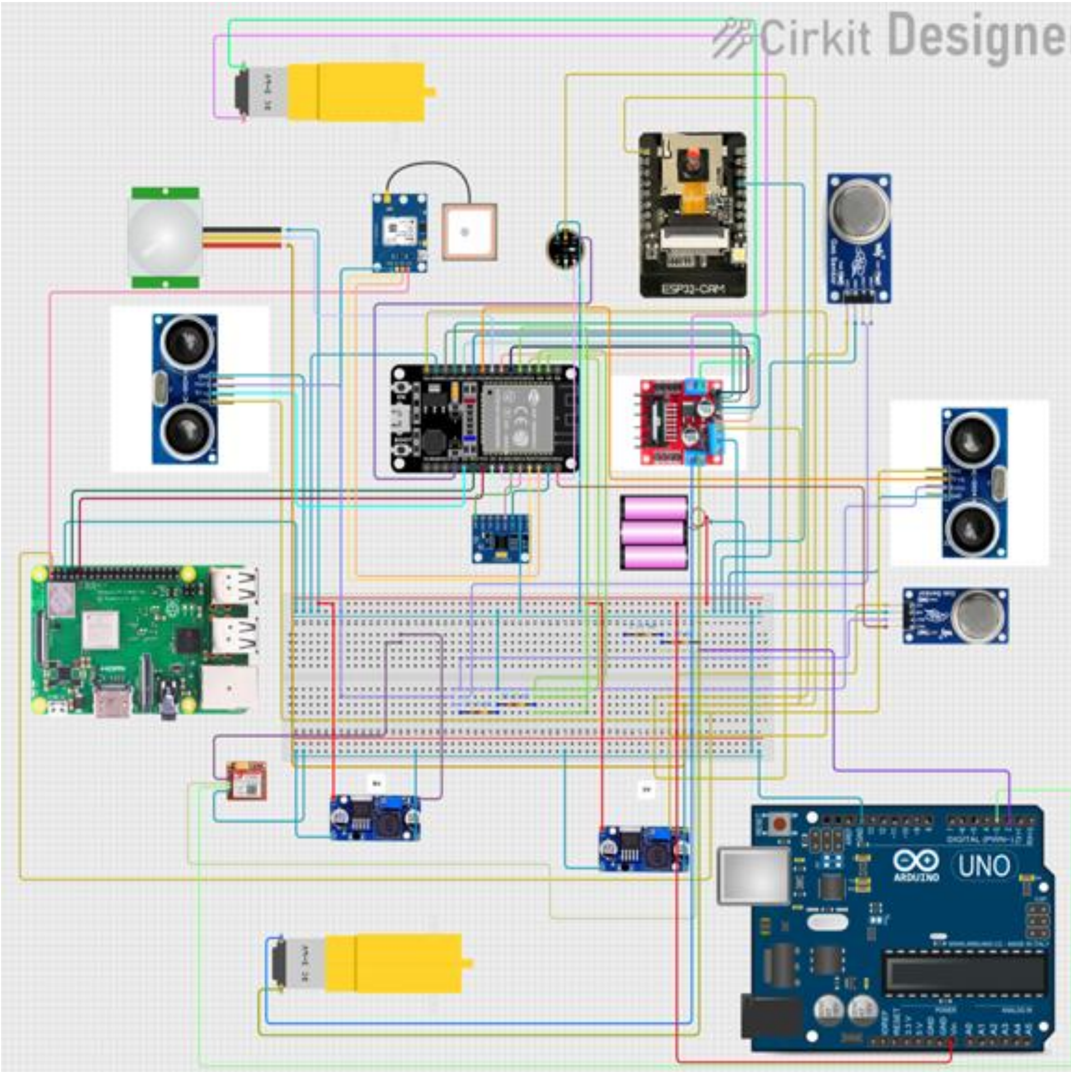

Figure 1: The created circuit diagram of the autonomous car using cirkitdesigner software

The environmental sensing is performed with MQ2 (flammable gases) and MQ135 (air quality) for hazard detection. Localization and navigation are carried out using a GPS NEO-6M connected to the ESP32, Ultrasonic sensors for obstacle avoidance and an IMU MPU6050 for stability. The Arduino Uno was dedicated exclusively to the SIM800L GSM module, that enabling reliable wireless communication with the user. Sensor data were transmitted to the Raspberry Pi, where AI inference and sensor fusion were performed. Reinforcement learning algorithms guided autonomous navigation in simulated disaster terrains. A web-based dashboard was developed with a Flask backend and a responsive frontend using HTML, CSS and JavaScript supported by an SQLite database for real-time monitoring, enabling survivor localization, hazard alerts, and vehicle tracking.

The experimental workflow was executed sequentially to ensure robust system development. Hardware integration and calibration were first performed for all modules, including ultrasonic sensors for obstacle detection, MPU6050 IMU for stability, INMP441 microphone for audio, MQ2 and MQ135 gas sensors, and GPS NEO-6M and GSM SIM800L for localization and communication. AI models were then trained on disaster-relevant datasets: YOLOv5 for visual human detection, a CNN for thermal image classification, and an audio classifier for survivor voice recognition. Outputs from these models and sensors were integrated via a fusion and reinforcement learning framework on the Raspberry Pi 3B+, enabling adaptive navigation and autonomous decision-making. The prototype was tested in controlled indoor environments simulating obstacles, debris, and varying gas concentrations. Finally, a Flask-based web dashboard with HTML/CSS/JavaScript frontend and SQLite database provided real-time monitoring through WebSocket communication, delivering continuous visualization of detections, hazards, and survivor locations. The stepwise experimental workflow for system development is illustrated in Figure 2, which outlines the processes from hardware calibration through AI training, sensor fusion, and final prototype validation.

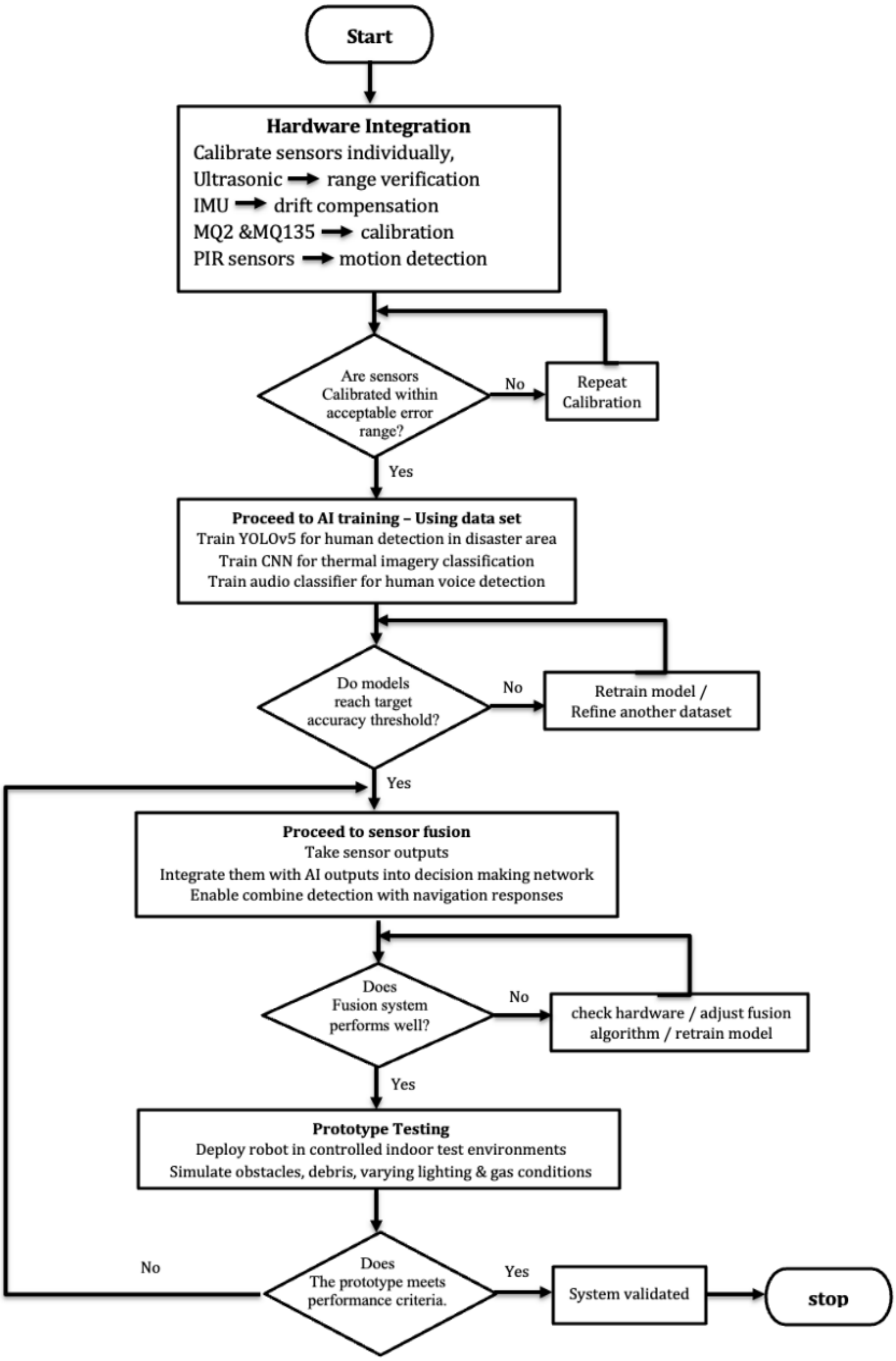

Figure 2: The experimental process of testing of the constructed system

Preliminary trials validated subsystem performance before integration. Ultrasonic sensors reliably detected obstacles up to 3 meters away, while the MPU6050 confirmed stable orientation tracking. Gas sensors (MQ2 and MQ135) responded accurately to low concentrations of LPG and CO_2 under laboratory conditions. The CNN classifier consistently distinguished human thermal signatures, and the audio classifier detected human voices with acceptable accuracy under controlled noise. These results guided the configuration of sensor fusion thresholds and decision-making rules for the prototype from the Figure 3.

Using a dataset of 1,467 training images and 571 validation images, the thermal CNN achieved the performance metrics reported in Table 1, demonstrating robust generalization on unseen thermal data.

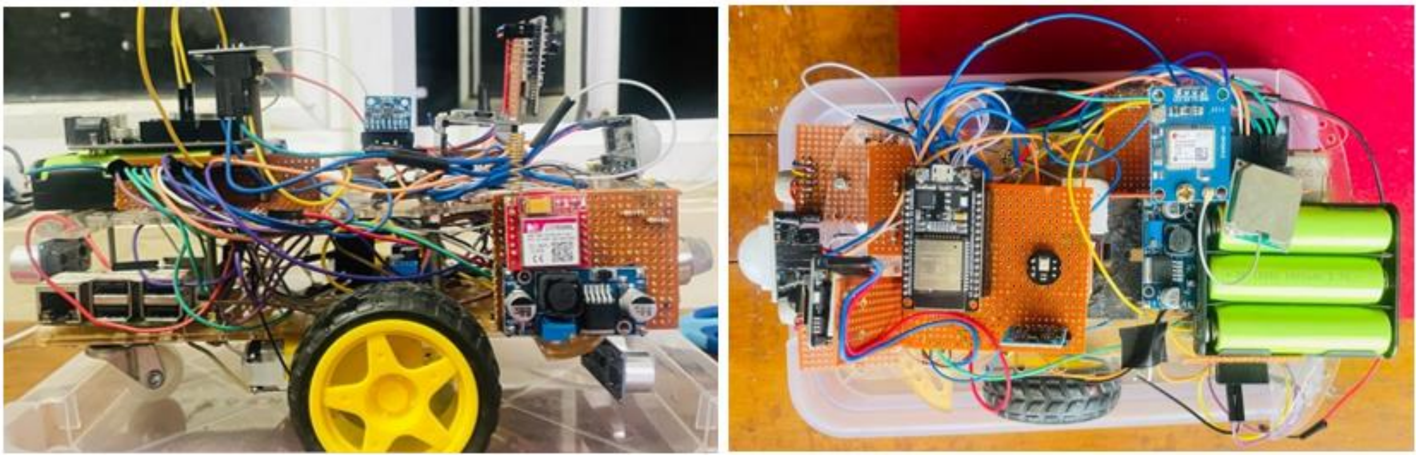

Figure 3: The up and side views of the experimental setup for the system

The CNN was optimized using the Binary Cross Entropy (BCE) loss function, defined in Equation (1), to train the thermal image classifier with y_i representing the ground truth label (human or non-human) and $\hat{y}_i$ representing the projected probability. Accurately separating human heat signatures from background regions is encouraged by this loss function which penalizes incorrect classifications.

Equation 1

$$L_{CE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\ \log(1 - \hat{y}_i)\,]$$

Where:

N = The number of samples

$y_i \in \{0,1\}$ = The ground truth label

$\hat{y}_i$ = The predicted probability from the CNN

At the current stage, navigation is achieved through ultrasonic-based obstacle avoidance and reinforcement learning–guided decision rules. Advanced SLAM-based mapping is not implemented in this prototype and is identified as a key direction for future enhancement.

RESULTS AND DISCUSSION

The autonomous robot prototype was successfully implemented and tested in controlled environments simulating disaster-like conditions. The ultrasonic sensors consistently detected obstacles up to 3 m with an average accuracy of 94%. The MPU6050 IMU further stabilized navigation by compensating for sudden tilts.

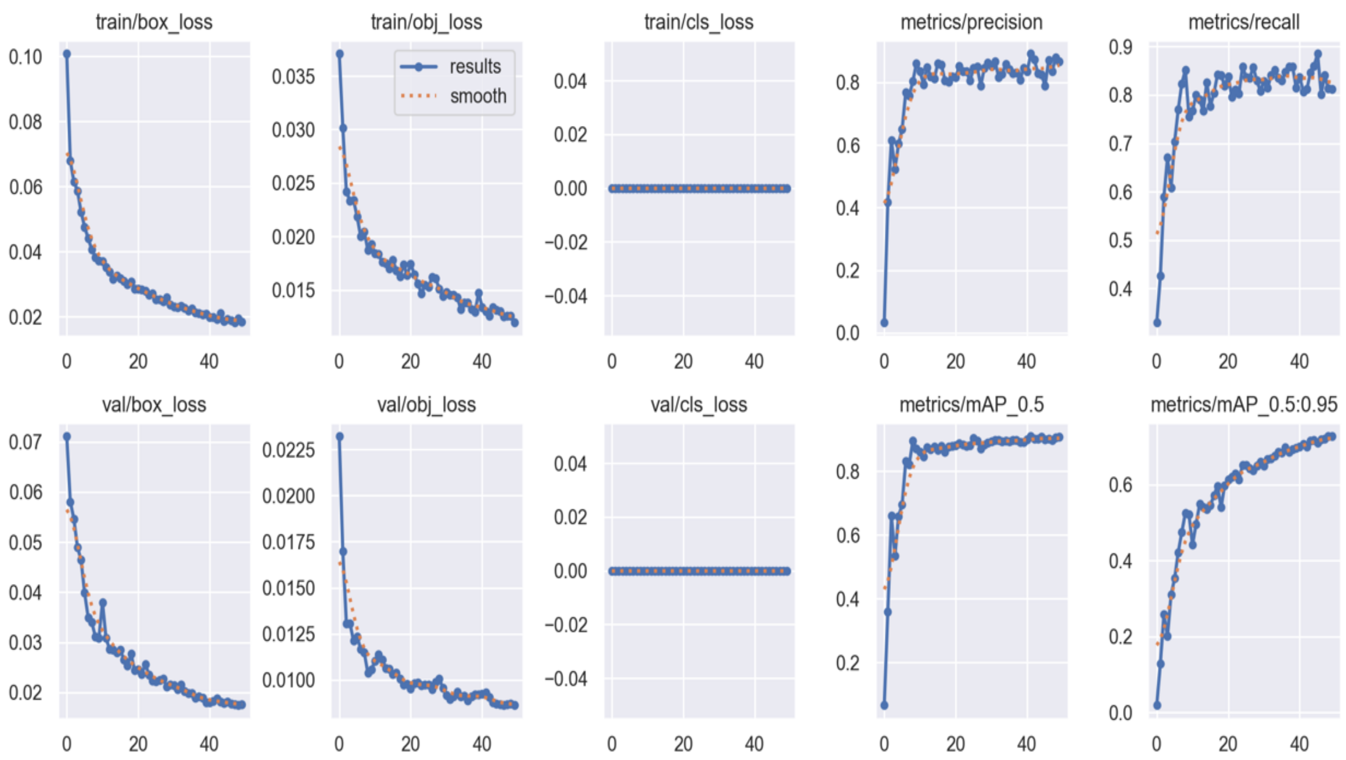

Figure 4: Training and validation performance of the YOLOv5 model. The plots show training/validation losses (box, objectness, classification) and evaluation metrics (precision, recall, mAP@0.5, mAP@0.5:0.95) over 50 epochs. The decreasing losses and increasing metrics indicate good model convergence and generalization

The training process of YOLOv5 is illustrated in Figure 4, where the losses decrease smoothly, and performance metrics converge to stable high values, demonstrating effective learning. Figure 5 illustrates that the model achieved high accuracy in distinguishing speech from non-speech.

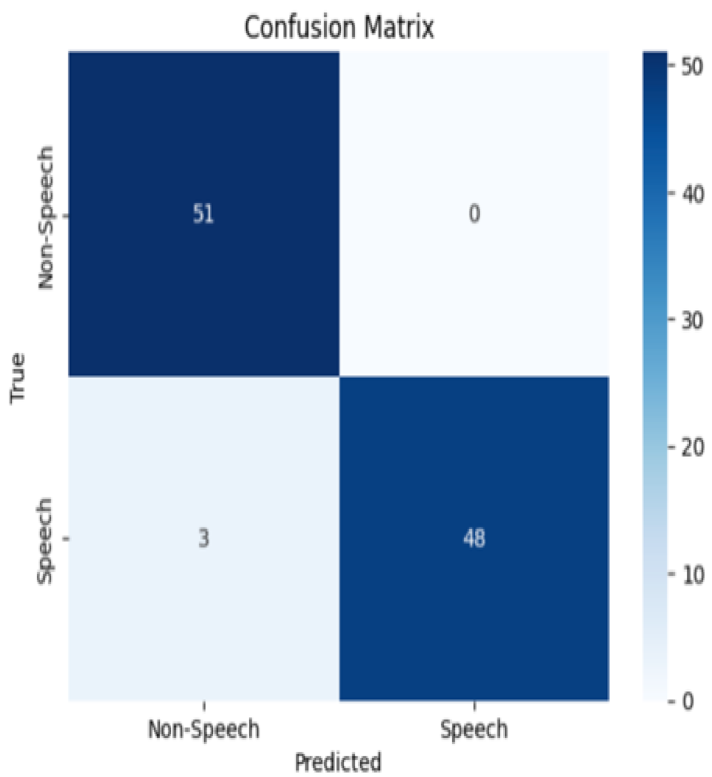

Figure 5: Confusion matrix of the speech vs. non-speech classification. The model correctly classified 51 non-speech and 48 speech samples, with only 3 speech samples misclassified.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue XII December 2025

The MQ2 and MQ135 sensors responded effectively to varying concentrations of LPG and CO_2, providing hazard alerts when thresholds exceeded safe ranges. This demonstrated the system's ability to identify toxic or flammable environments. The web-based dashboard successfully provided real-time updates using WebSocket communication, enabling continuous display of survivor localization, sensor readings, and hazard alerts without page reload. Response latency averaged less than 1 second between the robot and base station, supporting practical field deployment [6]. As illustrated in Figure 6, the training and validation losses decrease steadily while precision, recall, and mAP improve over the course of 50 epochs.

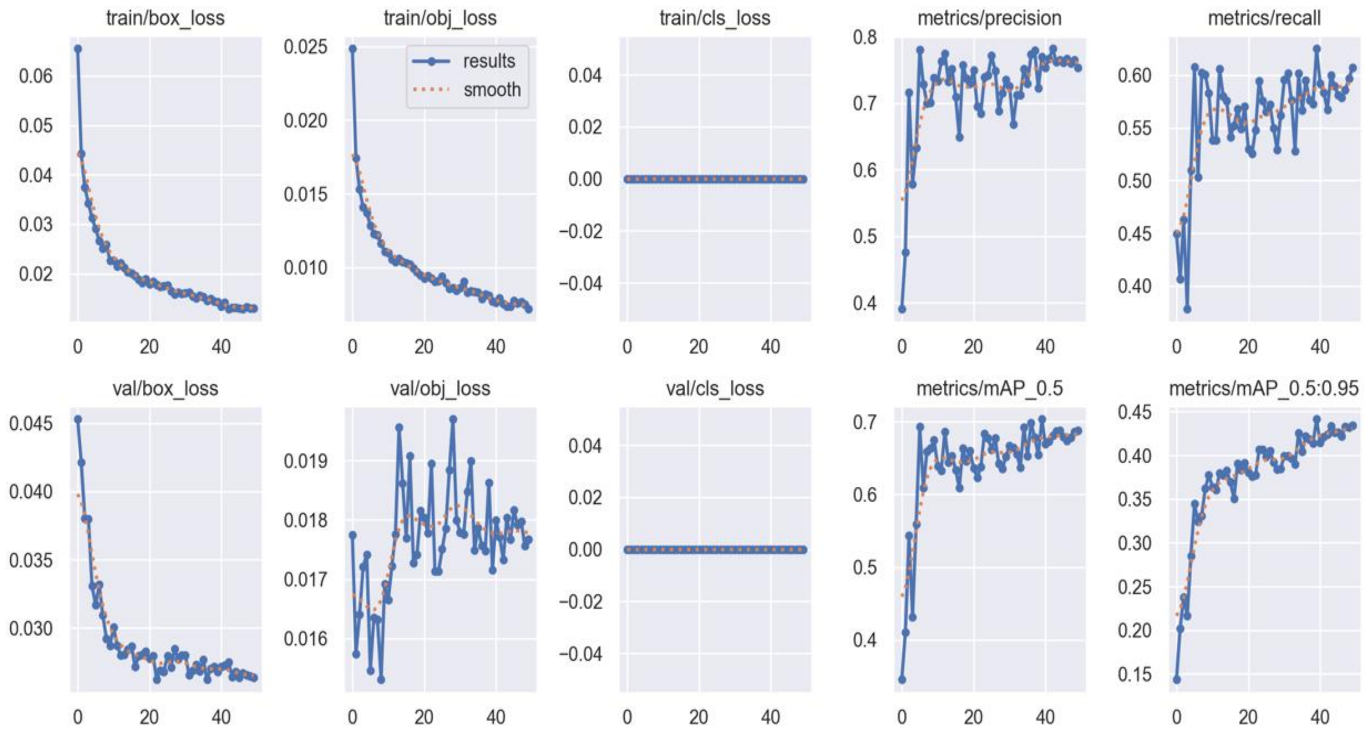

Figure 6: Training and validation metrics of the CNN-based object detector on the thermal dataset, showing loss convergence and improvements in precision, recall, and mAP over 50 epochs.

The performance of the thermal CNN model evaluated on the thermal dataset is summarized in table 1. The model achieved an accuracy of approximately 93-94% with balanced precision and recall, indicating reliable human detection in thermal imagery. The CNN was optimized using the binary cross entropy loss function Equation 1, with minimizes the classification error between predicted and ground truth labels. The convergence behavior of this loss function is evident in the performance results summarized in Table 1, where the model achieves high precision and recall on the thermal dataset.

Table 1: CNN performance on thermal dataset

Metric	Value
Accuracy	~0.93-0.94
Precision	~0.91
Recall	~0.92
F1-score	~0.915
Training Loss	~0.07-0.08
Validation Loss	~0.09-0.10

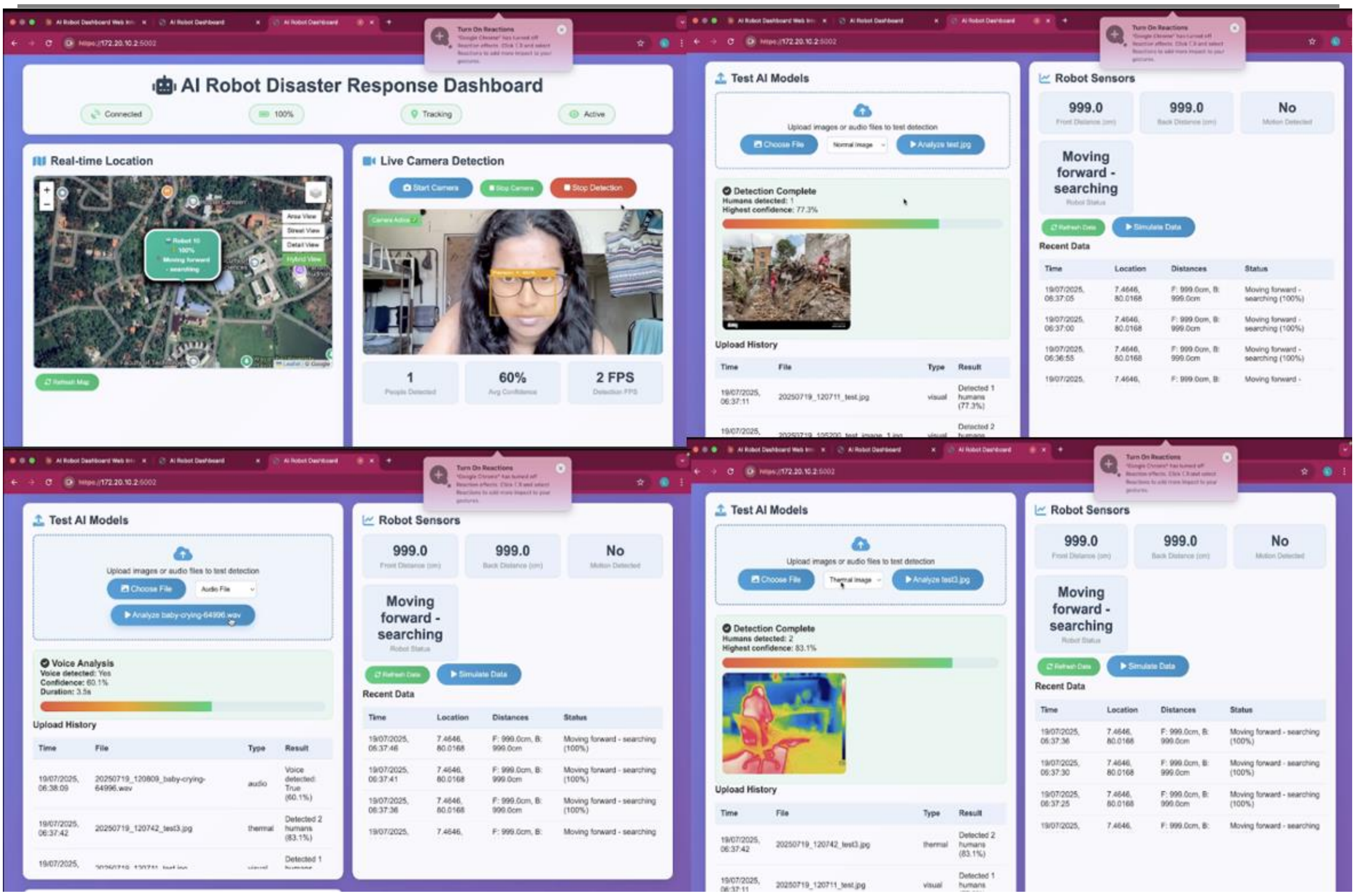

Figure 7: Four images of the web dashboard

Figure 7 shows the AI Robot Disaster Response Dashboard, which provides real-time tracking of the robot's location, live camera feed, AI model results, and sensor data. It allows for monitoring the robot's progress and performance during disaster response operations.

LIMITATIONS

Despite the promising results, the proposed system has several limitations. The integration of multiple sensors and AI modules increases system complexity and may affect long-term reliability in harsh disaster environments. Additionally, the computational and power constraints of the Raspberry Pi 3B+ and ESP32 limit the simultaneous execution of multiple deep learning models. Performance may also degrade under extreme conditions such as dense smoke, high ambient noise, or severe occlusions. Furthermore, the navigation strategy is currently limited to obstacle avoidance and does not incorporate full SLAM-based mapping, which may restrict scalability in large disaster sites.

CONCLUSION

The prototype effectively demonstrated autonomous navigation and victim detection capabilities in controlled, simulated disaster environments. Ultrasonic sensors provided accurate and reliable obstacle avoidance, ensuring smooth navigation around debris and barriers. The YOLOv5 visual detection model achieved robust human detection, confirming its suitability for disaster victim identification. The CNN for thermal image classification reliably distinguished human heat signatures from the background, augmenting detection robustness under low-visibility conditions. Meanwhile, the audio classifier trained on human voice samples successfully identified potential survivors through acoustic cues, complementing the visual and thermal detection modalities [6]. Gas sensors (MQ2 and MQ135) effectively detected hazardous concentrations, enabling timely hazard alerts. The integration of these sensor inputs through a sensor fusion and reinforcement learning framework on the Raspberry Pi 3B+ facilitated real-time decision-making and adaptive navigation. Furthermore, real-time monitoring via the web-based dashboard empowered rescue operators with continuous updates on survivor localization and environmental hazards, enhancing situational awareness. The experimental evaluation using the physical prototype demonstrates an average visual detection accuracy of 89%, thermal classification accuracy of 90% and audio detection accuracy of 71% with an obstacle avoidance success rate of 93% and reliable dashboard data transmission in 95% of test scenarios. These results prove that the proposed system is a promising, affordable and rapidly deployable solution for disaster response applications. Collectively, these results validate the feasibility of a cost-effective, AI-driven robotic platform for disaster response applications, demonstrating consistent, multi-modal survivor detection and autonomous navigation [7].

ACKNOWLEDGEMENT

Authors would be grateful to the lecturers and staff at the Department of Electronics, Faculty of Applied Sciences, Wayamba University of Sri Lanka for their guidance and support, and laboratory facilities**.**

REFERENCES

1. Anyfantis, A., Silis, A., & Blionas, S. (2021). A low-cost, mobile e-nose system with an effective user interface for real-time victim localization and hazard detection in USaR operations. Measurement: Sensors, 16, Article 100049. https://doi.org/10.1016/j.measen.2021.100049
2. Biggie, H., & McGuire, S. (2022, March 23). Heterogeneous ground-air autonomous vehicle networking in austere environments: Practical implementation of a mesh network in the DARPA Subterranean Challenge. ResearchGate. https://www.researchgate.net/publication/359450481
3. Leong, W. L., Cao, J., & Teo, R. (2024). Dynamic decentralized 3D urban coverage and patrol with UAVs. arXiv. https://arxiv.org/abs/2406.09828
4. Mazhar, O., Babuska, R., & Kober, J. (2021). GEM: Glare or gloom, I can still see you – End-to-end multi-modal object detection. IEEE Robotics and Automation Letters, 6(4), 6321–6328. https://doi.org/10.1109/LRA.2021.3093871
5. Mărieș, M., & Tătar, M. O. (2025). Design and simulation of mobile robots operating within networked architectures tailored for emergency situations. Preprints. https://doi.org/10.20944/preprints202504.1747.v1
6. Merkle, N., et al. (2023). Drones4Good: Supporting disaster relief through remote sensing and AI. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW 2023) (pp. 3772–3776). IEEE. https://doi.org/10.1109/ICCVW60793.2023.00407
7. Panagopoulos, D., Perrusquia, A., & Guo, W. (2024). Selective exploration and information gathering in search and rescue using hierarchical learning guided by natural language input. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC 2024) (pp. 1175–1180). IEEE. https://doi.org/10.1109/SMC54092.2024.10831125