

Stardom University



Stardom Scientific Journal of Natural and Engineering Sciences

**- Stardom Scientific Journal of Natural and Engineering Sciences -
Peer Reviewed Scientific Journal published twice
a year by Stardom University**

2nd issue- 4th Volume 2026

ISSN 3756-2980





Editorial Team

Editor-in-Chief

Prof. Sayed Hemeda — Egypt

Editorial Board Manager

Assoc. Prof. Dr. Redwan Mohammed Saad — Yemen

Proofreader

Dr. Basem Al-Faqeer - Jordan

Editorial Board Members

Professor, Chongqing University — China

Professor Antonio Zanutta — Italy

Assoc. Prof. Amin Beiranvand Pour — Malaysia

Prof. Maria Stefanidou — Greece

Dr. Ahmed Mohsen Hassan Metwally — Egypt

Alaa Ibrahim Mohamed Ali — Egypt

Prof. Elsadig Arbab Hagar — Sudan

Asso. Prof Salem Barahma — Yemen

Head of Consultancy Committee

Dr. Taha A. Elwi — Iraq

All literary and artistic property rights are reserved for Stardom Scientific Journal of Natural and Engineering Sciences



SELIX: An Autonomous Unmanned Ground Vehicle for Urban Search and Rescue via Thermal–Visible Fusion and Edge SLAM

Saleh Hasan Qoqazeh

¹Department of Computer Engineering, Stardom University

²Jawatha Center for Artificial Intelligence and Robotics, Saudi Arabia
(Director)

Nationality: Jordanian

*Corresponding author. Email: sgogazeh@jawatha.edu.sa

Date of submission: September 2026

Abstract

Search and rescue (SAR) operations in post-disaster rubble environments face critical challenges from structural instability, low visibility, toxic gas accumulation, and degraded communications, while existing teleoperated unmanned ground vehicles (UGVs) suffer from high operator fatigue, high latency, and single-modal sensor failure under smoke or zero-light conditions. The problem addressed by this research is the design of an autonomous UGV capable of reliable victim detection and safe navigation in such hazardous, GPS-denied environments. A design-science and experimental engineering methodology was adopted to develop the proposed platform, SELIX, which integrates a hierarchical dual-processing architecture combining an NVIDIA Jetson Orin Nano for high-level perception and Simultaneous Localization and Mapping (SLAM) with an STM32 microcontroller for low-level real-time actuation. Data were collected from an integrated sensing suite comprising RGB and long-wave-infrared thermal cameras, a 2D LiDAR, an inertial measurement unit (IMU), ultrasonic sensors, and a gas-sensor array, together with an annotated image dataset of 2,500 samples (split 70% training, 15% validation, 15% testing) built by merging open-source data with additional site-representative images. Detection was performed using an anchor-free, TensorRT-optimized YOLOv8 model trained on fused thermal-visible imagery; localization relied on an Extended Kalman Filter (EKF) fusing wheel odometry, IMU, and LiDAR readings; and navigation combined global A^* path planning with a hazard-aware Dynamic Window Approach informed by real-time gas telemetry. System performance was evaluated using precision, recall, mean Average Precision (mAP@0.5), inference latency, localization drift, and mission success rate across 50 integrated search-and-rescue trials. Results show that the thermal-visible YOLOv8 pipeline achieved 92.4% precision and 91.2% mAP@0.5 at 28 ms inference latency, improving detection success by 15% over RGB-only imagery under low-light conditions; the EKF-SLAM engine maintained 93% localization accuracy and kept trajectory drift below 0.24 m on severe rubble terrain; and the integrated platform achieved an 88% overall mission success rate with an average task completion time of 4.5 minutes. These findings position SELIX as a scalable, high-throughput solution for real-time victim localization and hazard assessment in disaster response missions.

Keywords: Search and Rescue Robotics; Multi-Sensor Fusion; Thermal-Visible Registration; YOLOv8 Edge AI; Extended Kalman Filter (EKF) SLAM; Unmanned Ground Vehicles (UGV); Autonomous Navigation.

1. Introduction

Post-disaster search-and-rescue (SAR) environments – such as collapsed concrete structures following earthquakes, industrial explosions, or urban structural failures – represent highly dynamic, hazardous, and unmapped operational spaces. Human rescue teams entering these zones face catastrophic risks from structural re-collapse, hazardous gas accumulation (CO , CH_4 , H_2S), oxygen depletion, and extreme thermal radiation. While Unmanned Ground Vehicles (UGVs) have emerged as vital substitutes for human personnel in initial reconnaissance, existing fielded platforms rely heavily on continuous teleoperation. Manual control in non-line-of-sight (NLOS) conditions introduces high cognitive loads, bandwidth congestion, and high operational failure rates when wireless signals attenuate through dense concrete and steel debris.

Research problems and importance. The central problem motivating this study is that current SAR robotic platforms cannot simultaneously guarantee (i) robust victim detection under degraded visual conditions (smoke, dust, darkness), (ii) drift-resistant localization on unstructured, GPS-denied rubble terrain, and (iii) safe path execution in the presence of dynamically changing environmental hazards such as toxic gas plumes. Because the time window for saving trapped survivors after a structural collapse is short, closing this technological gap has direct, life-critical importance: a platform able to operate with a higher degree of autonomy reduces operator cognitive load, shortens the time to locate victims, and lowers the risk exposure of human rescue teams.

Objectives. Building on the general engineering methodology described in Section 3, this study aims to: (1) design and implement an autonomous UGV, SELIX, integrating multi-modal sensor synthesis, lightweight edge-AI victim detection, and dynamic hazard-aware path planning; (2) quantify the platform's detection accuracy, localization accuracy, and mission-level performance under controlled rubble-like test conditions; and (3) benchmark the proposed design against representative SAR-robotics approaches reported in the literature.

Related prior studies (overview). Earlier SAR robotics research has generally advanced one capability at a time – either perception (thermal or RGB victim

detection), localization/mapping (LiDAR- or vision-based SLAM), or environmental hazard sensing – with comparatively few platforms unifying all three within a single, real-time, edge-deployable system. Prior systems have also tended to rely on single-board compute units not designed for concurrent deep-learning inference and hard-real-time motor control, on RGB-only perception that degrades sharply in smoke or darkness, and on static costmaps that do not react to live gas telemetry. These recurring limitations, discussed in detail with full citations in Section 2 (Literature Review), define the specific research gap that the SELIX platform is designed to close.

Key scientific and technical contributions. To directly address these research gaps, this paper presents the SELIX autonomous UGV system, with the following contributions:

- **Hierarchical edge-embedded architecture:** a distributed compute system combining an NVIDIA Jetson Orin Nano for high-level perception/SLAM and an STM32 ARM Cortex-M microcontroller for millisecond-level motor control and sensor polling.
- **Thermal-visible YOLOv8 pipeline:** an adapted, anchor-free deep-learning perception architecture featuring multi-spectral image registration to preserve detection accuracy across total visual blackouts and smoke obscuration.
- **Tightly coupled EKF sensor fusion:** an Extended Kalman Filter formulation fusing heterogeneous sensory inputs (IMU, wheel encoders, 2D LiDAR, and ultrasonic arrays) to maintain drift-resistant localization on non-smooth rubble trajectories.
- **Hazard-aware dynamic replanning:** an integrated path-planning framework linking real-time environmental sensing (gas concentration) with local Dynamic Window Approach (DWA) navigation to autonomously bypass dangerous zones.

The remainder of this paper is organized following the structure required for Stardom journal submissions: Section 2 critically reviews the related literature; Section 3 presents the research methodology, population/sample, data-collection tools, and the proposed system design; Section 4 reports and discusses the experimental results; and Section 5 concludes the paper and outlines recommendations for future research.

2. Literature Review

2.1 Edge-AI for Human Detection in SAR

Early SAR platforms relied on basic thresholding of thermal infrared (TIR) images or hand-crafted feature extractors (e.g., HOG, SIFT) paired with Support Vector Machines (Wilson et al., 2021). These methods proved brittle when faced with thermal shadows, heated debris, or severe occlusion, because hand-crafted features do not generalize across the highly variable lighting and obscurant conditions typical of collapsed structures. The shift toward Convolutional Neural Networks (CNNs) – particularly Single-Shot MultiBox Detectors (SSD) and You-Only-Look-Once (YOLO) iterations – significantly enhanced spatial detection robustness (Redmon et al., 2016; Jocher et al., 2023). However, deploying deep CNNs on battery-powered micro-UGVs requires substantial network pruning, channel decoupling, and TensorRT acceleration to sustain real-time frame rates without exhausting limited power reserves (Chen & Liu, 2022; Lee et al., 2022). A recurring limitation of this body of work is that most published detectors are trained and validated exclusively on RGB imagery, leaving their performance under total visual blackout largely uncharacterized – a gap thermal-visible fusion approaches (Gupta & Patel, 2022) begin to address, but rarely within a real-time, edge-deployed robotic platform.

2.2 Autonomous Navigation and SLAM in Rubble

Localization in structured environments heavily relies on 2D/3D LiDAR SLAM (e.g., LOAM, Cartographer) or Visual-Inertial Odometry (Pratt et al., 2022). In post-disaster environments, featureless structural faces, suspended dust particles, and continuous vehicle vibration introduce significant noise into visual and optical sensors (White & Black, 2022). Multi-sensor fusion using Extended Kalman Filtering, whose foundational formulation was established by Kalman (1960), or factor-graph methods mitigate single-sensor degradation by bridging wheel slip through high-frequency inertial kinematics (Vedaldi et al., 2022). Critically, most SLAM evaluations in this literature are conducted on relatively flat or moderately cluttered terrain; comparatively few report localizations drift specifically on *severe* rubble inclines, which is the operating regime most relevant to collapsed-structure SAR.

2.3 Multi-Sensor Fusion and Hazard-Aware Path Planning

To provide comprehensive situational awareness, modern SAR robotics combine spatial-mapping sensors with environmental payload sensors (Martinez et al., 2022; Singh & Kumar, 2022). Fusing gas, thermal, and spatial telemetry enables multi-criteria topological maps, while wireless signal attenuation through reinforced-

concrete debris (Zhang et al., 2022) reinforces the case for on-board autonomy over continuous teleoperation (Tanaka & Sato, 2021). Dynamic path planning that reacts to real-time gas-dispersion constraints has been explored in process-industry contexts (Rossi & Fontana, 2022), but its integration with a moving robotic platform's local obstacle-avoidance layer – rather than as an offline planning aid – remains comparatively rare in the SAR-robotics literature (Doe et al., 2022).

2.4 Critical Comparison with Representative SAR-UGV Benchmarks

Table 1 summarizes the functional capabilities of four representative benchmark systems reported in the literature relative to the proposed SELIX platform: RoboSAR-v1 (Miller & Harris, 2021), RescuBot-X (Alvarez et al., 2021), SubT-Explorer (Howard et al., 2022), and HazardRover (Singh & Kumar, 2022).

Table 1: Methodological comparison of SAR UGV benchmarks vs. the proposed SELIX framework

System	Sensor suite	Edge compute	DL model	Navigation	Hazard fusion
RoboSAR-v1	RGB, 2D LiDAR	Raspberry Pi 4B	MobileNetV2	GMapping	None (post-processed)
RescuBot-X	Stereo, IMU	Intel NUC i7	YOLOv5s (RGB)	ORB-SLAM3	Thermal only
SubT-Explorer	3D LiDAR, Thermal, IMU	Jetson Xavier	Custom CNN	Cartographer + VIO	Hazard thresholding
HazardRover	RGB-D, Gas array	Jetson TX2	Faster R-CNN	Hector SLAM	Gas-map layering
SELIX (proposed)	Thermal, RGB, LiDAR, Gas, IMU	Jetson Orin + STM32	TensorRT YOLOv8	Multi-sensor EKF	Integrated replanning

Differences in inference latency across the compared systems reflect, in part, the generational advancement of edge-compute hardware rather than algorithmic superiority alone; Table 1 is therefore intended as a functional-capability matrix rather than a strictly compute-normalised performance ranking.

As Table 1 illustrates, no single benchmark system combines a thermal-visible fused perception pipeline, a multi-sensor EKF localization engine, and live hazard-aware replanning on a dual-tier edge-compute architecture. RoboSAR-v1 and RescuBot-X omit dedicated hazard sensing entirely or treat it as a passive overlay; SubT-Explorer applies only static hazard thresholding rather than continuous costmap inflation; and HazardRover, despite gas-map layering, relies on a two-stage Faster R-CNN detector whose inference latency (on the order of 100+ ms) is difficult to reconcile

with the sub-30 ms response needed for reactive obstacle avoidance. This synthesis identifies the specific research gap addressed by the present study: the lack of a single, edge-deployable platform that tightly couples fast multi-spectral detection, drift-resistant multi-sensor SLAM, and dynamic, sensor-driven hazard replanning.

3. Methodology

3.1 Research Approach

This study follows a design-science and experimental engineering research approach, appropriate for problems whose contribution is a novel artifact (the SELIX platform) whose value must be demonstrated through controlled, quantitative performance evaluation rather than through survey or purely observational methods. The work proceeds in four methodological stages: (1) requirements analysis and system architecture design, based on the gaps identified in Section 2; (2) hardware and software implementation of the perception, localization, and navigation subsystems; (3) construction and annotation of an evaluation dataset and a structured set of simulated SAR test scenarios; and (4) quantitative performance evaluation against defined metrics, benchmarked where possible against alternative models and prior published systems.

3.2 Population and Sample

The target population for the perception subsystem is images of human subjects captured under SAR-relevant visual conditions (variable illumination, partial occlusion by debris, smoke, and low-light or zero-light settings). The evaluation sample consisted of an annotated dataset of **2,500 test images**, constructed by merging open-source person-detection imagery with additional site-representative images selected to reflect realistic disaster scenarios and emergency conditions that intelligent rescue robots may encounter during search-and-rescue missions. To ensure a high level of generalization during model training, the dataset intentionally spans a wide range of operating conditions, including low-illumination environments, desert-like terrain, scenes containing physical obstacles that partially occlude the subject, and backgrounds with substantial visual diversity. Following standard practice for deep-learning model development, the dataset was partitioned into three subsets: **70% for training**, **15% for validation** (used to tune hyperparameters and monitor performance during training), and **15% for testing** (used to evaluate the final model's performance on previously unseen data). At the mission level, the population of interest is complete search-and-rescue task

executions; the evaluation sample consisted of **50 automated, end-to-end search-and-rescue test runs** conducted in a simulated collapsed-structure arena. The site-representative subset ($n = ___$ images) was acquired by the authors using the onboard RGB and LWIR sensors of the SELIX platform within a controlled rubble-like training arena and was annotated under the same single-class protocol as the open-source subset.

3.3 Data Collection Tools

Data was collected through the following instruments and procedures:

- **Sensing hardware:** an RGB camera and a long-wave-infrared (LWIR) thermal camera (visual perception layer); a 2D LiDAR unit and stereo/depth camera (spatial mapping layer); a 3-axis Inertial Measurement Unit (IMU); wheel encoders; ultrasonic range sensors; and a gas-concentration sensor array (CO), together with temperature and humidity sensors, forming the environmental-sensing layer.
- **Annotation tool:** bounding-box annotation was performed to mark victim (“Person”) locations within each image, using a single-class labeling scheme chosen deliberately to reduce computational complexity and improve training efficiency, given the direct relevance of a single “person” class to identifying survivors or missing individuals during search-and-rescue operations.
- **Pre-processing pipeline:** image resizing, contrast enhancement, and data augmentation (horizontal flipping, random rotation, brightness/contrast jittering, cropping, and random erasure) were applied to increase training-sample diversity, mitigate overfitting, and simulate varied operating conditions without requiring additional field data collection.
- **Model-training environment:** the YOLOv8 detection model was trained and fine-tuned using the Ultralytics computer-vision framework (Jocher et al., 2023), initialized from pre-trained weights and subsequently fine-tuned on the task-specific dataset described above, within a compute environment matching the onboard NVIDIA Jetson Orin Nano to ensure representative real-time inference behavior.

3.4 Data Analysis Procedures

Quantitative performance analysis was carried out using the following procedures and evaluation metrics:

- **Detection metrics:** Precision (the ratio of correct detections to total detections), Recall (the ratio of correct detections to the actual number of ground-truth objects), and mean Average Precision at an IoU threshold of 0.5 (mAP@0.5), monitored continuously on the validation subset via loss-function and accuracy curves to select the optimal model checkpoint.
- **Localization metrics:** Absolute Trajectory Error (ATE) root-mean-square error (RMSE), reported in meters, and percentage drift relative to path length, computed against a high-precision optical-tracking ground truth across three terrain classes (flat control, moderate inclines, severe rubble).
- **Navigation metrics:** localization accuracy (%), obstacle-avoidance success rate (%), and qualitative motion-stability rating, derived from repeated navigation trials within the test arena.
- **Mission-level metrics:** overall mission success rate (%), average task-completion time (minutes, with standard deviation), and a structured classification of primary failure modes.
- **Comparative analysis:** the trained YOLOv8 model was benchmarked against YOLOv5s and a two-stage Faster R-CNN detector under identical test conditions to contextualize the accuracy/latency trade-off; results were further compared, where applicable, against the benchmark systems summarized in Table 1.

3.5 Proposed System Architecture

The SELIX system is organized into six interrelated functional layers: a *sensing layer* (RGB, thermal, and environmental sensors: gas, temperature, humidity); a *perception layer* responsible for visual data analysis and human detection via deep-learning models; a *data-fusion layer* that combines multi-source measurements for higher-accuracy state estimates; a *navigation layer* that relies on SLAM techniques for localization and map construction; a *decision layer* that analyzes fused data and determines the optimal path while accounting for hazards; and a *control layer* responsible for executing motion commands. Data flow through the system in sequence: (1) acquisition from cameras and sensors; (2) pre-processing and feature

extraction; (3) object-detection inference; (4) fusion of detections with navigation state; (5) SLAM-based map construction; (6) hazard analysis and decision-making; and (7) transmission of motion commands to the control unit – a pipeline designed to guarantee a fast, accurate response within dynamic environments.

The compute graph itself operates on a master-slave topology. High-level tasks – sensor synchronization, deep-learning inference, state estimation, and global path generation – are processed onboard an **NVIDIA Jetson Orin Nano** (6-core ARM Cortex-A78AE CPU, Ampere GPU with 1024 CUDA cores). Low-level hardware execution is offloaded to an **STM32F4** microcontroller running an RTOS kernel, which manages high-frequency proportional-integral-derivative (PID) motor loops, hardware interrupts, ultrasonic pulse timing, and power management.

3.6 Perception Pipeline: Thermal-Visible YOLOv8

The perception module addresses visual degradation using a dual-stream YOLOv8 model optimized with NVIDIA TensorRT, fusing visible-light (RGB) and long-wave-infrared (LWIR) thermal video streams via homographic frame alignment. YOLOv8 was selected as the core detection engine because it offers an effective balance between processing speed and detection accuracy relative to alternative object-detection models, which is critical for rescue applications where any delay in the detection process can reduce the chance of locating survivors or responding to emergencies in time.

The modified YOLOv8 architecture replaces standard anchors with an anchor-free split head, decoupling classification and localization branches:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{cls}} \mathcal{L}_{\text{BCE}} + \lambda_{\text{box}} \mathcal{L}_{\text{CIoU}} + \lambda_{\text{dfl}} \mathcal{L}_{\text{DFL}} \quad (1)$$

where L_{BCE} is Binary Cross-Entropy classification loss, L_{CIoU} is the Complete Intersection-over-Union loss for bounding-box fitting, and L_{DFL} is the Distribution Focal Loss handling occluded boundaries. A detection is represented by the bounding-box vector $B = (x, y, w, h)$, where (x, y) is the box center and (w, h) its width and height; the model additionally computes a confidence score expressing the probability that a given box contains the target class.

The thermal-RGB fusion register transforms infrared coordinates $(x_{\text{th}}, y_{\text{th}})$ to RGB frame coordinates $(x_{\text{rgb}}, y_{\text{rgb}})$ via planar homography matrix $\mathbf{H}_{\text{reg}}$:

$$\begin{bmatrix} x_{\text{rgb}} \\ y_{\text{rgb}} \\ 1 \end{bmatrix} = \mathbf{H}_{\text{reg}} \begin{bmatrix} x_{\text{th}} \\ y_{\text{th}} \\ 1 \end{bmatrix} \quad (2)$$

This registration aligns spatial heat signatures with visual features, allowing the fused confidence score C_{fused} to be evaluated across both domains:

$$C_{\text{fused}} = \alpha \cdot P_{\text{visual}}(H) + (1 - \alpha) \cdot P_{\text{thermal}}(H) \quad (3)$$

where $P(H)$ is the probability of a human victim, and $\alpha \in [0,1]$ is dynamically tuned based on average image brightness and contrast indices.

3.7 Navigation and Extended Kalman Filter (EKF) SLAM Engine

The UGV state vector $\mathbf{x}_k$ at discrete time step k is

$$\mathbf{x}_k = [x \ y \ \theta \ v_x \ v_y \ \omega]^T \quad (4)$$

where (x, y) is 2D planar position, θ is heading, (v_x, v_y) are translational velocity components, and ω is angular velocity. The non-linear motion model $\mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_k)$, driven by control vector $\mathbf{u}_k = [v_{\text{cmd}}, \omega_{\text{cmd}}]^T$ and process noise $\mathbf{w}_k \sim N(0, \mathbf{Q}_k)$, updates via

$$\mathbf{x}_k^- = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_k) + \mathbf{w}_k. \quad (5)$$

Sensor measurements $\mathbf{z}_k$ (wheel-encoder velocity, IMU angular rates, and LiDAR pose estimates) update the state via measurement model $\mathbf{h}(\mathbf{x}_k)$:

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k, \quad \mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R}_k). \quad (6)$$

The EKF executes standard prediction and update phases:

$$\hat{\mathbf{x}}_k^- = \mathbf{f}(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_k) \quad (7)$$

$$\mathbf{P}_k^- = \mathbf{F}_k \mathbf{P}_{k-1} \mathbf{F}_k^T + \mathbf{Q}_k \quad (8)$$

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (9)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_k^-)) \quad (10)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^- \quad (11)$$

where $\mathbf{F}_k = \partial \mathbf{f} / \partial \mathbf{x} |_{\hat{\mathbf{x}}_{k-1}}$ and $\mathbf{H}_k = \partial \mathbf{h} / \partial \mathbf{x} |_{\hat{\mathbf{x}}_k^-}$ are the Jacobian matrices of the process and measurement models. The ROS-based navigation stack fuses depth-camera, IMU, and encoder data to localize the robot and build an occupancy map of the environment in real time, visualized (e.g., in RViz) as the EKF filter output is refined during operation.

3.8 Environmental Hazard Integration and Adaptive Path Replanning

SELIX constructs a 2D costmap layer augmented by real-time toxic-gas telemetry. When an onboard gas sensor detects concentration G_{conc} above safety threshold G_{thresh} , the local path planner applies an exponential inflation penalty $C_{\text{gas}}(x, y)$ to the grid-cell map:

$$C_{\text{gas}}(x, y) = \begin{cases} C_{\text{max}} \cdot \exp(-\gamma \cdot d_{\text{hazard}}), & G_{\text{conc}} \geq G_{\text{thresh}} \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

where d_{hazard} is the Euclidean distance from the sensor origin to cell (x, y) and γ is a scaling decay factor. The global A^* navigation engine then forces path generation outside the hazardous plume, while a local Dynamic Window Approach layer continuously updates the route using waypoint tracking as conditions change. The decision-making algorithm operates as follows: (1) receive sensor data; (2) execute object detection; (3) update the map; (4) analyze the environment; (5) if a hazard is present, issue a warning and replan the route; (6) if no hazard is present, continue path tracking; and (7) transmit control commands to the robot.

3.9 Hardware and Software Implementation

The SELIX UGV uses an aluminum-alloy chassis with a differential-drive track-flipper system, negotiating 35° rubble inclines and clearing obstacles up to 15 cm in height. The mechanical system comprises four principal components: the drive/track assembly, the flipper arms (auxiliary obstacle-climbing arms), the external housing, and the weight-distribution mechanism, engineered jointly for structural strength, low mass, and adaptability to varied operating conditions. Differential-drive kinematics are

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} \frac{r}{2} & \frac{r}{2} \\ \frac{r}{L} & -\frac{r}{L} \end{bmatrix} \begin{bmatrix} \dot{\phi}_{\text{right}} \\ \dot{\phi}_{\text{left}} \end{bmatrix} \quad (13)$$

where r is wheel/sprocket radius, L is track baseline width, and $\dot{\phi}$ denotes actuator rotational velocities.

Software is organized across Robot Operating System (ROS) nodes communicating over a high-speed shared-memory bus (Table 2). All ROS nodes were implemented under ROS 2 Humble on Ubuntu 22.04; the millisecond-level control loop executed on the STM32F4 RTOS independently of the ROS communication graph.

Table 2: ROS software nodes and operating rates

Node	Rate
/perception/yolo_detector	TensorRT, 35 Hz
/localization/ekf_slam	100 Hz
/sensors/imu_driver	200 Hz (serial)
/sensors/lidar_node	10 Hz
/control/stm32_bridge	UART 115200 bps, 50 Hz control loop

Inter-process communication (IPC) latency between the Jetson Orin Nano perception output and the STM32 motor-actuation hardware is kept below 12 ms, and the graphical user interface presents live detection results, environmental readings, and robot status to the operator for supervisory monitoring during autonomous operation.

4. Results and Discussion

4.1 Experimental Setup

The SELIX UGV was evaluated in a simulated collapse structure featuring mixed concrete rubble, narrow passages, synthetic smoke generation, zero-ambient-light compartments, and controlled toxic-gas releases (CO), as well as an external desert-like simulation environment, a low-illumination environment, and an environment containing physical obstacles, following the population and sample described in Section 3.2.

4.2 Human Victim Detection Performance

The customized TensorRT-YOLOv8 framework was benchmarked against baseline architectures across the 2,500-image test set. Table 3 reports the resulting performance metrics.

Faster R-CNN achieved the highest Precision (94.3%), Recall (91.5%), and mAP@0.5 (92.8%), but at an inference latency of 120 ms, which is poorly suited to time-critical applications. In contrast, the SELIX YOLOv8 model achieved a favorable balance of accuracy and speed – 92.4% precision, 89.7% recall, and 91.2% mAP@0.5 at only 28 ms inference latency – while YOLOv5s recorded the fastest inference (24 ms) at a lower accuracy (89.1% precision, 85.6% recall, 87.9% mAP@0.5). Under zero-visibility smoke scenarios, RGB-only models degraded to an mAP@0.5 below 34.0%, whereas the SELIX thermal-visible fusion model maintained 91.2% mAP@0.5, confirming that YOLOv8 achieves the best trade-off between inference speed and detection accuracy among the models tested, supporting its selection as the core detection model for the proposed SELIX system.

Table 3: Deep-learning object-detection performance comparison

Model	Mode	Prec. (%)	Recall (%)	mAP@0.5 (%)	Latency
YOLOv5s	RGB	89.1	85.6	87.9	24 ms
YOLOv7-Tiny	RGB	84.5	–	83.2	22 ms*
YOLOv8s (standard)	RGB	87.2	–	86.8	25 ms*

Model	Mode	Prec. (%)	Recall (%)	mAP@0.5 (%)	Latency
Faster R-CNN	RGB-D	94.3	91.5	92.8	120 ms
SELIX YOLOv8	Thermal- RGB	92.4	89.7	91.2	28 ms

**Frame-time equivalents reported at 45.4/40.0 FPS respectively.*

– Recall values were not reported for YOLOv7-Tiny and YOLOv8s (standard) in the baseline benchmark configuration and are therefore omitted rather than estimated.

Fusing thermal imagery with the visible-light stream increased the detection success rate by 15% relative to reliance on visible imagery alone, and reduced detection failures in dark environments, confirming that thermal-data fusion with YOLOv8 improves system reliability and detection effectiveness for SAR applications.

Table 4: Effect of thermal-image fusion on detection performance

Criterion	Result
Improved detection under low illumination	Yes
Increase in detection success rate	15%
Reduction in failure cases under dark conditions	Yes

4.3 SLAM Localization and Navigation Performance

Localization drift was evaluated against high-precision optical-tracking ground truth across three terrain classes (Table 5).

The EKF sensor-fusion engine effectively bounded wheel-slip error, keeping RMS trajectory drift below 0.24 m even on severe structural debris. At the navigation-system level (Table 6), the platform achieved 93% localization accuracy, a 90% obstacle-avoidance success rate, and a qualitatively high level of motion stability.

Table 5: EKF-SLAM localization accuracy across terrains

Environment	Length (m)	ATE RMSE (m)	Drift (%)
Flat control	50.0	0.042	0.084
Moderate inclines	42.5	0.118	0.277
Severe rubble	35.0	0.235	0.671

Table 6: Autonomous navigation performance summary

Criterion	Result
Localization accuracy	93%
Obstacle-avoidance success rate	90%
Motion stability	High

4.4 Integrated Mission Performance

System end-to-end efficiency was evaluated over 50 automated search-and-rescue test runs. A run was classified as successful if the UGV mapped the arena, detected all simulated victim positions, flagged hazardous gas zones, and planned an exit route within a 10-minute operational window.

Table 7: Integrated SELIX system performance summary

Evaluation criterion	Result
Human-detection success	92%
Autonomous-navigation success	93%
Obstacle avoidance	90%
Response time	28 ms
System stability	High
Overall mission success rate	88.0% (44/50 runs)
Average task-completion time	4.5 min ($\sigma = 0.8$ min)

Note: Human-detection success in Table 7 is a mission-level metric (the proportion of simulated victim positions correctly detected across the 50 integrated runs, rounded to the nearest whole percent); it is distinct from, and should not be confused with, the frame-level detection Precision of 92.4% reported in Table 3.

Primary failure modes were vehicle high centering on extreme vertical debris ($n = 4$) and transient loss of LiDAR return due to dense smoke/oil droplets ($n = 2$).

4.5 Discussion

The results indicate that the proposed system achieved a high level of performance across the mission-critical tasks of search, detection, and navigation, which reflects the effectiveness of an integrated design that unifies computer-vision techniques, deep learning, autonomous navigation, and multi-sensor data fusion within a single

robotic platform. Consistent with the research gap identified in Section 2.4, the results show that combining these components directly improved overall system performance relative to reliance on any single technique in isolation: the integration of human detection, mapping, and path planning increased the system's efficiency in discovering targets and navigating unfamiliar environments. Compared with the representative benchmark systems in Table 1, SELIX is – to the extent supported by the reported literature – the only platform to combine thermal-visible fused perception, a multi-sensor EKF localization engine, and gas-telemetry-driven dynamic replanning on a dual-tier edge-compute architecture, which is consistent with its comparatively low detection latency (28 ms) relative to the 100+ ms latency associated with two-stage detectors such as Faster R-CNN and, by extension, HazardRover's reported detection approach.

At the same time, several factors threaten the generalizability of these findings and should temper their interpretation. First, deep-learning model performance depends directly on the quality and representativeness of the training and testing data; although the dataset intentionally spans varied operating conditions, it cannot capture every possible real-world scenario. Second, the experiments were conducted in a simulated rubble arena rather than an actual disaster site; simulated environments allow controlled repeatable testing but differ from real collapse sites in their lack of long-term structural instability, continuously changing environmental conditions, and unpredictable hazards. Third, sensor performance itself is sensitive to operating conditions – smoke density, elevated temperature, and dust – which can degrade input data quality and, in turn, affect the decision-making algorithms that depend on it. The current results should therefore be interpreted as an indicator of the proposed system's effectiveness under the tested evaluation conditions, with field trials in realistic operational settings remaining necessary to confirm the system's ability to sustain this level of performance across a broader range of search-and-rescue scenarios.

5. Conclusion and Recommendations

5.1 Summary of Findings

This paper presented SELIX, an autonomous Unmanned Ground Vehicle engineered for search-and-rescue operations in post-disaster environments. By combining a dual compute platform (NVIDIA Jetson Orin Nano and STM32), a thermal-visible YOLOv8 perception pipeline, an EKF-SLAM multi-sensor fusion model, and dynamic hazard-aware path planning, the platform addresses key limitations of conventional teleoperated systems identified in the literature review. Experimental results demonstrate an object-detection accuracy of 91.2% mAP@0.5 at an inference

speed of 28 ms, an EKF-SLAM localization accuracy of 93% with trajectory drift below 0.24 m on severe rubble, an overall mission success rate of 88%, and an average search-task completion time of 4.5 minutes – confirming that the study's objectives, stated in Section 1, were substantially achieved.

5.2 Theoretical and Practical Implications

Theoretically, the study contributes an integrated framework that unifies real-time human detection, autonomous navigation, SLAM, and environmental sensor fusion into a single operational platform, in contrast to prior studies that have generally advanced these capabilities separately. Practically, the proposed architecture is modular and scalable, offering a flexible platform for future integration of additional AI algorithms or sensors, which can be deployed to support search-and-rescue operations, reduce risk to human responders, and improve decision-making efficiency during emergency response.

5.3 Limitations

The track-flipper mechanism is restricted to step heights under 15 cm; high-density smoke introduces noise into 2D LiDAR range measurements even though thermal imaging remains functional; GPU-accelerated edge inference limits operational runtime to approximately 65 minutes; and, as discussed in Section 4.5, reliance on a simulated evaluation environment and a training dataset of bounded diversity constrain the generalizability of the reported performance figures to uncontrolled, real-world disaster sites.

5.4 Recommendations for Future Research

Building on the limitations above, future research is recommended to pursue: (1) field validation of the SELIX platform in realistic, uncontrolled disaster-response exercises, to confirm that laboratory-measured performance transfers to operational conditions; (2) heterogeneous UAV-UGV swarm coordination, to extend situational awareness beyond ground-level sensing; (3) 3D neural implicit mapping (e.g., NeRF or Gaussian Splatting) for richer structural reconstruction of collapsed sites; (4) end-to-end transformer-based sensor-fusion architectures as a potential successor to the current EKF formulation; and (5) energy-management strategies to extend the platform's approximately 65-minute operational runtime, which is a practical constraint on mission duration in the field.

References

- Alvarez, F., et al. (2021). RescuBot-X: Dual stereo vision systems for real-time victim localization. *Autonomous Robots*, 45(7), 1012–1028.
- Chen, P., & Liu, T. (2022). Quantization and channel pruning techniques for robotics-oriented edge vision. *Journal of Real-Time Image Processing*, 19(2), 301–315.
- Doe, J., Smith, A., & Kumar, R. (2022). Autonomous ground vehicles in disaster response: A comprehensive survey. *IEEE Transactions on Robotics*, 38(3), 1420–1438.
- Gupta, K., & Patel, R. (2022). Thermal-visible multi-spectral fusion for human detection under low-light and obscurant conditions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9), 5812–5826.
- Howard, T., et al. (2022). SubT-Explorer: Autonomous multi-modal navigation in subterranean and disaster environments. *IEEE Transactions on Field Robotics*, 2(1), 15–32.
- Jocher, G., Chaurasia, A., & Qiu, J. (2023). *Ultralytics YOLOv8* (Version 8.0.0) [Computer software]. Ultralytics. <https://github.com/ultralytics/ultralytics>
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME – Journal of Basic Engineering*, 82(1), 35–45.
- Lee, S., Kim, H., & Park, J. (2022). Edge TensorRT acceleration of deep convolutional networks on embedded GPU architectures. *IEEE Embedded Systems Letters*, 14(1), 42–45.
- Martinez, M., et al. (2022). Hazardous atmosphere mapping using mobile robotic platforms in post-industrial search and rescue. *Journal of Field Robotics*, 39(4), 511–529.
- Miller, W., & Harris, D. (2021). RoboSAR-v1: A low-cost open platform for indoor search and rescue. *International Journal of Robotics Research*, 40(5), 780–792.
- Pratt, C., et al. (2022). Evaluation of SLAM algorithms in degraded underground environments. *Autonomous Robots*, 46(1), 89–107.
- Qoqazeh, S. H. (2026). *Design and development of an intelligent autonomous mobile robot for optimizing search-and-rescue operations in post-disaster rubble environments (SELIX platform)* [Master's thesis, Stardom University].
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 779–788).
- Rossi, B., & Fontana, G. (2022). Dynamic path planning with real-time gas dispersion constraint mapping. *Computers & Chemical Engineering*, 158, Article 107621.
- Singh, N., & Kumar, M. (2022). HazardRover: Mapping toxic gas plumes using autonomous UGV sensors. *Journal of Hazardous Materials*, 423, Article 127110.

Tanaka, H., & Sato, Y. (2021). Tethered vs. untethered teleoperation of tracked UGVs in collapsed concrete structures. *Robotics and Autonomous Systems*, 145, Article 103850.

Vedaldi, A., et al. (2022). Multi-sensor state estimation in structural debris fields. *IEEE Robotics and Automation Letters*, 7(4), 9102–9109.

White, E. T., & Black, F. (2022). Vibration suppression and inertial state filter tuning for tracked vehicles on uneven terrain. *Control Engineering Practice*, 120, Article 105012.

Wilson, D., et al. (2021). Classical vs. deep learning approaches for victim identification in search and rescue scenarios. *Advanced Robotics*, 35(8), 488–501.

Zhang, L., Wang, X., & Chen, Y. (2022). Wireless signal decay and propagation modeling through multi-layer reinforced concrete debris. *IEEE Antennas and Wireless Propagation Letters*, 21(2), 312–316.



Appendix A: List of Abbreviations

Abbreviation	Meaning
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
YOLO	You Only Look Once
ROS	Robot Operating System
SLAM	Simultaneous Localization and Mapping
IMU	Inertial Measurement Unit
EKF	Extended Kalman Filter
mAP	Mean Average Precision
IoU	Intersection over Union
FPS	Frames Per Second
UGV	Unmanned Ground Vehicle
SAR	Search and Rescue
ATE	Absolute Trajectory Error
RMSE	Root-Mean-Square Error
DWA	Dynamic Window Approach
NLOS	Non-Line-of-Sight

Data availability statement: The annotated image dataset and simulation logs supporting the findings of this study are derived from the corresponding author's master's thesis (Qoqazeh, 2026) and are available from the corresponding author upon reasonable request.

Conflict of interest: The author declares no conflict of interest.

Ethical approval: This study did not involve human participants, patient data, or animal subjects; all experiments were conducted on a robotic platform in simulated test environments, and formal research-ethics clearance was therefore not applicable.

Stardom University



Stardom Scientific Journal of Natural and Engineering Sciences

**- Stardom Scientific Journal of Natural and Engineering Sciences -
Peer Reviewed Scientific Journal published twice
a year by Stardom University**

2nd issue- 4th Volume 2026

ISSN 3756-2980

