Autonomous Port Area Safety Inspection Using a Multimodal Sensor-Integrated UGV

Jen-Hao Liu[1], Wei-Ping Hung[2], Chun-Ying Lee[2], Jacob J. Lin[1], and Jen-Yu Han[1]

[1]Department of Civil Engineering, National Taiwan University, Taiwan
[2]Ministry of Transportation and Communications, Institute of Transportation, Taiwan

r13521709@ntu.edu.tw, welma@mail.ihmt.gov.tw, jimmylee@mail.ihmt.gov.tw, jacoblin@ntu.edu.tw, jyhan@ntu.edu.tw

Abstract -

Traditional port inspections are highly dependent on manual labor and static sensors, leading to high operational costs and fragmented data. These limitations hinder effective data integration and advanced analysis, creating a critical need for an automated inspection framework to improve safety and efficiency. This paper proposes an automated safety inspection system that uses Unmanned Ground Vehicles (UGVs). The system implements a multimodal data integration platform that synchronizes heterogeneous sensor streams by aligning timeframes and unifying sampling frequencies. Data transmission is facilitated via a HyperText Transfer Protocol (HTTP)-based real-time streaming architecture, enabling a centralized interface for remote waypoint planning and mission execution. To ensure operational safety, UGVs incorporate dynamic obstacle avoidance and an object detection model to identify critical assets. The system also performs real-time thermal monitoring, triggering automated alarms when temperature thresholds are exceeded. Field tests demonstrate that the UGV achieves precise autonomous navigation in outdoor port environments, with a mean waypoint navigating error within 0.032 meters. The proposed framework offers a centralized and accessible solution for safety inspections in extensive port environments, effectively reducing operational complexity while enhancing overall efficiency and personnel safety.

Keywords -

Unmanned Ground Vehicles (UGVs); Port Safety Inspection; Multimodal Data Integration; Autonomous Navigation; Real-time Monitoring

1 Introduction

Safety inspections are a critical component of seamless port operations. Given the high risks associated with industrial assets such as transformer boxes and heavy lifting equipment, there is a growing necessity for systematic and autonomous inspection frameworks to replace traditional, labor-intensive methods, as robotic inspection systems have been shown to greatly improve operational efficiency and provide more effective workflows compared to conventional approaches [1]. However, the expansive and complex nature of modern ports poses significant challenges to traditional monitoring frameworks. Currently, port safety relies primarily on two methods: manual labor or fixed sensor networks. Manual inspections are inherently problematic; Not only are they labor-intensive, but also highly susceptible to variations in personnel skill, physical fatigue, and environmental hazards [2]. Such inconsistencies often lead to qualitative and subjective assessments, resulting in delayed anomaly reporting and a lack of high-quality, continuous data required for long-term safety and structural health analysis. While fixed sensors provide a degree of automation, their static nature results in a limited detection range, creating "blind spots" and producing fragmented data that is difficult to synthesize into a coherent situational overview.

The scattered distribution of key assets within a port further complicates real-time data integration. Comprehensive monitoring requires diverse sensing technologies: thermal imaging for heat and anomaly detection, LiDAR for spatial awareness and mapping, and gas detectors for environmental hazards [3]. However, these multimodal sensors operate with heterogeneous data formats and varying sampling frequencies, presenting significant synchronization challenges [4]. Without a centralized system to align timeframes and unify transmission protocols, the gathered data remains isolated, underscoring the need for a smarter, scalable UGV-based inspection solution [5].

To address these limitations, UGVs offer a transformative approach to port safety by providing mobile, flexible sensing that can cover an entire port's infrastructure. Automated navigation enables precise, repeatable inspection routes ensuring consistent data collection across every critical node. Deploying UGVs in outdoor industrial settings requires sophisticated obstacle avoidance and robust GPS-guided positioning, while map-based interfaces lower the operational barrier so port staff can manage robotic fleets without specialized technical expertise.

This paper proposes an automated safety inspection system centered on UGVs, designed to enhance the resilience and efficiency of port management. This study makes the following contributions to address practical deployment challenges specific to port safety inspection environments:

1. **GPS-to-SLAM Coordinate Bridging**: A mathematically explicit coordinate transformation pipeline that projects GPS waypoints from geographic coordinates into the Gmapping local map frame via UTM projection and heading-offset rotation, enabling autonomous outdoor waypoint navigation without pre-built maps. This bridging mechanism between global GPS reference and local SLAM-generated occupancy grids addresses a practical integration challenge that is rarely formalized in prior inspection robot literature.

2. **Targeted Multimodal Thermal Fusion**: A fusion strategy that maps YOLOv8-generated RGB detection bounding boxes onto spatially corresponding thermal image regions to extract asset-specific temperatures, enabling infrastructure-level anomaly monitoring and reducing false positives caused by ambient thermal background noise.

3. **End-to-End Field-Validated Port Inspection Framework**: An integrated UGV-based inspection system validated in a real outdoor port environment, directly addressing the operational gaps specific to port infrastructure management: fragmented sensor data, heavy reliance on manual patrol labor, and the absence of an accessible centralized monitoring interface for non-specialist operators.

The remainder of this paper is organized as follows: Section 2 reviews related work in industrial inspection automation, multimodal sensor fusion, and robot navigation. Section 3 details the proposed method, including hardware configuration, ROS-based data collection, real-time alert mechanisms, and waypoint navigation. Section 4 describes the experimental setup and case study design. Section 5 presents the experimental results and discusses the system's limitations. Finally, Section 6 summarizes the entire paper and looks forward to future research directions.

2 Related Work

2.1 UGV-based Inspection and Navigation

Traditional port inspection relies on manual patrols or stationary sensor networks, both of which suffer from high labor intensity, susceptibility to human error, limited spatial coverage, and fragmented data [6, 7, 8]. Gmapping has emerged as a practical SLAM benchmark for autonomous systems [9, 10], and integrating mobile robots with AI-based vision enables fully automated inspection paradigms [11]. However, most existing work addresses indoor logistics or general outdoor navigation without the targeted waypoint logic required for port asset inspection [12]. Indoor robots depend on pre-built maps, UAV-based systems face regulatory restrictions in active port zones, and manual inspection yields qualitative assessments difficult to integrate into long-term safety workflows [1, 13, 2]. The proposed UGV framework directly addresses these gaps through continuous, remotely supervised inspection with a single operator interface.

2.2 Data Integration and Real-time Monitoring

Automated inspection of critical assets such as transformer boxes requires integrating LiDAR, thermal, and visual sensing, yet multisensor fusion faces significant hurdles including data imperfection and synchronization of heterogeneous signals [4, 3]. Wu et al. [5] demonstrated robust LiDAR-visual-inertial fusion for high-fidelity positioning, while Page et al. [2] introduced modular "sensor bricks" to encapsulate sensing and communication. Despite these advances, no prior work delivers an end-to-end framework that closes the loop from raw multimodal acquisition to high-level decision support accessible to non-specialist port operators [13]. This research addresses that gap by integrating resilient multimodal sensing with a centralized remote monitoring platform tailored to port safety inspection.

3 Methodology

To realize the autonomous inspection using UGV, this research presents a method to navigate the UGV to perform a safety inspection and on-site data collection in unknown environments. As shown in Figure 1, the proposed method comprises three modules: coordinate navigation, multimodal data integration framework, and real-time warning system. The coordinate navigation module enables the UGV to perform a user-specified coordinate navigation under an unknown environment. The multimodal data integration framework aligns timeframes and unifies data frequencies from different sensors, enabling real-time multimodal data collection. The real-time warning system presents the collected data, so that the user, or inspector, can react to the warning in a short time.

3.1 Coordinate Navigation

To achieve autonomous inspection, this research designs a method for remotely assigning inspection waypoints to the UGV. This method comprises remote waypoint planning, coordinate conversion, Gmapping SLAM, and obstacle avoidance and navigation.

The remote waypoint planning module enables operators to assign inspection waypoints through the platform, with data transmission facilitated through a client-initiated HTTP polling architecture. To balance responsiveness

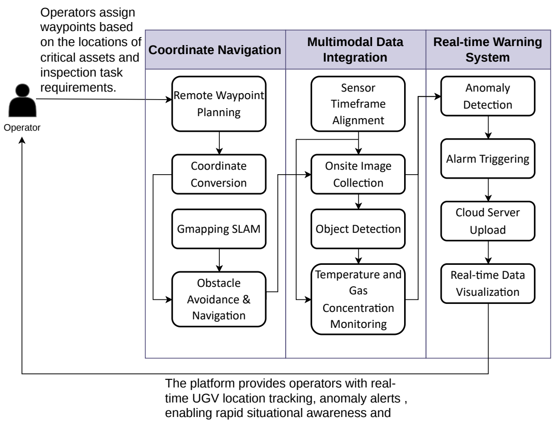

Figure 1. System workflow illustrating the integration of data collection, warning, and navigation modules.

with hardware efficiency, the system utilizes a 30-second polling interval for waypoint command updates. This interval is appropriate for port safety inspection for the following reasons: (1) inspection targets such as transformer boxes and oil tanks are fixed infrastructure assets whose positions do not change between or during missions, meaning pre-planned routes rarely require dynamic mid-mission modification; (2) the UGV's average traversal time between consecutive waypoints is on the order of several minutes, making a 30-second update cycle more than sufficient to capture any operator-initiated route adjustments; and (3) safety-critical data—including temperature readings, gas concentration levels, GPS coordinates, and on-site images—is transmitted immediately upon anomaly detection via a separate event-driven mechanism (Section 3.3), entirely independently of the polling cycle. The 30-second waypoint polling interval therefore does not constrain the system's responsiveness to safety hazards.

The coordinate conversion module transforms GPS waypoint coordinates into the local map frame generated by Gmapping, allowing the navigation planner to compute feasible paths within the occupancy grid. Initially, longitude and latitude coordinates are projected onto the Universal Transverse Mercator (UTM) system. This projection maps the ellipsoidal earth surface onto a 2D Cartesian plane, providing metric-based coordinates for precise local navigation. Let (E_0, N_0) represent the initial UTM position of the UGV at system startup. Its UTM coordinate displacement to the target waypoint (E_w, N_w) is:

$$\Delta E = E_w - E_0, \quad \Delta N = N_w - N_0 \tag{1}$$

Let ψ represent the heading offset between the UTM north axis and the local map y-axis, which is determined during initialization. The waypoint's position (x_m, y_m) in the local map coordinate system is:

$$\begin{bmatrix} x_m \\ y_m \end{bmatrix} = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} \Delta E \\ \Delta N \end{bmatrix} \tag{2}$$

The offset ψ is obtained by aligning the UGV to true north during initialization, and can then be recorded as the angular difference between the known north direction and the initial heading on the local map.

LiDAR collects laser scan data and transmits raw point cloud data to the edge computing unit. In the edge computing unit, the imported point cloud data can be used to create a 2D occupancy grid map through the Gmapping algorithm, a SLAM algorithm based on Rao-Blackwellized Particle Filter (RBPF) to generate a 2D occupancy grid map and robot poses from laser scan and odometry data.

The obstacle avoidance and navigation module uses path planning algorithms for autonomous waypoint navigation and dynamic obstacle avoidance. The system is divided into two planning modules for path planning. The global planner calculates the optimal path from the current position to the target waypoint based on a static occupancy grid map generated by Gmapping. The local planner continuously adjusts the trajectory by integrating real-time LiDAR data to avoid dynamic obstacles. All obstacles in the cost map are applied with an expansion radius of 0.5 meters to ensure safe passage during navigation. If the local planner cannot find a feasible path after multiple attempts, the current waypoint task is aborted and reported to the monitoring platform, requiring the operator to reassign new waypoints.

3.2 Multimodal Data Integration

The UGV inspection system proposed in this research integrates multiple sensors with different data formats and different sampling frequencies, such as LiDAR at 10 Hz and a depth camera at 30 Hz. To address the challenge of synchronizing multimodal data for further data processing and analysis, this research presents a data integration framework that aligns timeframes, unifies frequencies, and transmits data via HTTP protocols. Specifically, the framework employs the ROS `message_filters` package with an ApproximateTime synchronization policy to temporally align data streams from sensors operating at heterogeneous sampling frequencies. Each sensor node publishes timestamped messages to its dedicated ROS topic; the ApproximateTime synchronizer matches messages from different topics within a configurable time tolerance window, ensuring that multimodal data packets used for joint analysis—such as pairing RGB images with GPS coordinates for geolocated anomaly reporting—are temporally co-registered. The maximum synchronization latency introduced by this approach is bounded by the period of the lowest-frequency sensor (0.1 s for LiDAR at 10 Hz), which is well within the acceptable range for the quasi-static port inspection scenario targeted in this work. The framework is designed with a modular structure, making it simple and quick to add or remove sensors from the existing system. As shown in Figure 2, data received from various sensors is transmitted via USB to an edge computing unit, where the system aligns the timeframes and performs further data processing.

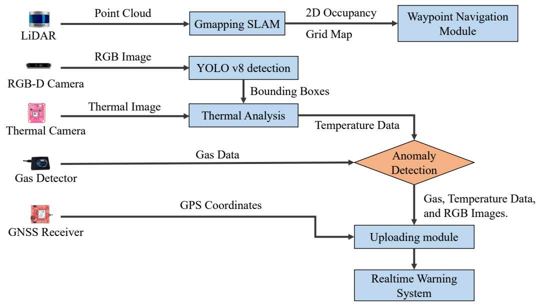

Figure 2. Software architecture for multimodal data integration, illustrating the data flow from sensors to edge computing unit for timeframe alignment, object detection, and anomaly monitoring.

The depth camera provides depth image and RGB image for the edge computing unit. In this study, we already have accurate depth information collected from LiDAR, so we only collect RGB images for the confirmation of actual on-site conditions.

The thermal camera captures infrared radiation emitted by objects and converts it into thermal images, which then provided to the edge computing unit. Temperature analysis performed on the edge computing unit includes both environment high-temperature anomaly detection and object high-temperature anomaly detection. The temperature alert thresholds for both can be adjusted according to the usage scenario. The YOLOv8s (small) model was employed for object detection in RGB images (11.2 M parameters, 28.6 GFLOPs), pre-trained on COCO 2017 and fine-tuned on a custom dataset of 293 annotated RGB images of transformer boxes collected on-site at the port. Data augmentation included HSV color-space, geometric (translation, scaling, flipping), and mosaic strategies. Training ran for 100 epochs (batch 16, input 640×640, SGD optimizer, lr_0=0.01, early stopping patience 50). The inference confidence threshold was set to 0.3, NMS IoU threshold 0.7. The detected bounding boxes were then mapped to corresponding thermal image regions to extract the highest and average temperatures for targeted high-temperature anomaly detection, reducing false positives from ambient thermal background noise. Gas detectors express flammable gas concentration in %LEL; an alarm threshold is set in accordance with ISO 10156 and IEC 60079 standards to provide early warning to operators.

3.3 Real-time Warning System

The alarm system uses the HTTP protocol via a 5G network to transmit sensor data from the edge computing unit to a cloud server. When an anomaly is detected, alarm information, along with relevant sensor readings, timestamps, and GPS coordinates, are immediately uploaded for real-time monitoring. Under normal circumstances, UGV's GPS coordinates are continuously uploaded at a frequency of 1 Hz to allow users to check the UGV inspection status on the platform. The anomaly data (including temperature data, gas data, timestamps, GPS coordinates, and images) are transmitted immediately upon detection of anomalies. If the anomaly persists, it will be transmitted at a frequency of approximately 0.2 Hz. When the network connection is interrupted, the edge computing unit will activate an offline storage mechanism to temporarily store the data locally and automatically synchronize the anomaly data upon reconnection.

4 Experimental Setup and Case Study

4.1 Hardware Configuration

This research employs the Clearpath Jackal UGV as a mobile robotic platform, loaded with a set of sensors and an onboard computer to perform autonomous outdoor inspection tasks. The Mini-ITX microcomputer (Intel Core i3-9100TE quad-core processor, 16 GB of DDR4 RAM, 250 GB SSD) is used to run the system. A Velodyne VLP-16 360° LiDAR scanner is connected to the microcomputer by ethernet cable to generate a 2D map describing the occupancy of the environment. The following sensors are connected to the microcomputer by USB: An Intel RealSense D455 depth camera, a FLIR Lepton FS thermal camera (through PureThermal 3 adapter board), an MPS003 gas detector (via USB-to-UART connector), and a high-precision u-blox ZED-F9P GNSS receiver. The first three are used for the data collection and potential hazard detection tasks, while the last is used for outdoor waypoint navigation tasks. A 5G router was also mounted on the UGV, providing a connection between the microcomputer and the external network. The connection between UGV, microcomputer, and sensors are highlighted in Figure 3.

4.2 ROS-based Data Integration and Streaming

This study uses ROS to implement the multimodal data integration framework described in Section 3. Each sensor publishes timestamped data to its corresponding ROS topic; the `message_filters` ApproximateTime policy

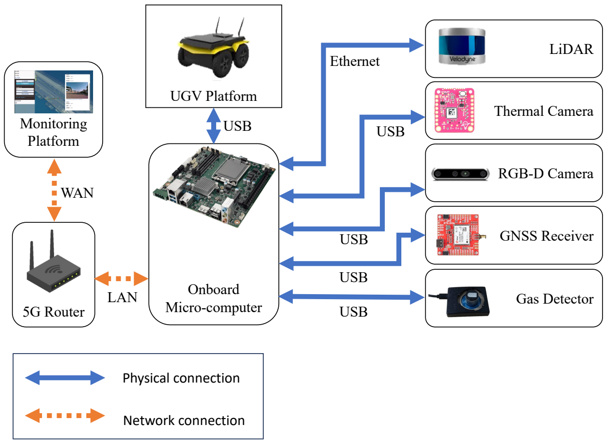

Figure 3. System connectivity diagram illustrating multimodal sensor integration and 5G data flow.

aligns heterogeneous data streams (e.g., LiDAR at 10 Hz and depth camera at 30 Hz) for joint processing and analysis. The *move_base* package handles autonomous waypoint navigation and dynamic obstacle avoidance.

4.3 Case Study

The primary focus is on the equipment and method design of an automated UGV outdoor safety inspection system. The system further integrates remote data presentation and route assignment platforms to enable operators to instantly get on-site information and react in advance when potential hazards occur. The GPS device can only perform precise localization tasks under open-sky conditions, and the four-wheeled UGV used in this study performs best for flat roads and large-scale environments. Considering these characteristics, this study selected the Taipei Port area as the experimental site, as this area has flat terrain and ample open space, as shown in Figure 4.

The transformer boxes are focused in the object high-temperature anomaly detection module. Because transformer boxes are critical infrastructure in the port area, situations such as short circuits could cause widespread damage and losses of personnel and facilities. The anomaly detection framework in this research encompasses both thermal and gas concentration monitoring. For temperature thresholds, a baseline of 45°C is established, as the sensors we used are rated for optimal operation below this limit. Consequently, this value serves as the high-temperature anomaly threshold. Furthermore, to identify critical overheating, an immediate alarm is triggered if any pixel within the thermal imagery exceeds 60°C. Regarding gas detection, an alarm threshold is set at 25% LEL as an early warning indicator for potential hazards, facilitating timely emergency protocols such as site evacuation or automated system shutdown. The on-site experiment aimed to verify the integrated capabilities of waypoint navigation, multi-sensor data acquisition, and port assets anomaly detection. An inspection route was assigned through a real-time monitoring platform. The assigned route covered key inspection targets, including transformer boxes and open roads.

Figure 4. Experimental site at Taipei Port.

5 Results

5.1 UGV Navigation Performance

The global path planner calculates viable paths for the UGV within the cost map generated by Gmapping, while the local planner dynamically adjusts the trajectory to avoid obstacles within an approximately 7-meter field of view. The system successfully bypassed all detected obstacles, including temporarily parked vehicles, electrical boxes, guardrails, and mooring posts, without any collisions occurring throughout the experiment. Position tracking analysis shows that, under normal conditions, the average deviation of the UGV from the planned path remains within 0.068 meters. In open clear-sky dock areas, GPS positioning accuracy is within 0.05 meters. When obstacles create narrow passages or dead ends, the path planner sometimes fails to identify feasible paths. After multiple failed attempts to recover (backwards, in-situ rotations), the navigation task is aborted, requiring the operator to re-designate the waypoint sequence.

Figure 5 presents the satellite map view of the navigation experiment conducted at Taipei Port. The yellow squares indicate the specified target waypoints assigned through the remote monitoring platform, while the blue circles represent the actual positions where the UGV arrived. The close alignment between target points and actual positions across all nine waypoints demonstrates the high precision of the GPS-based navigation system in open outdoor environments.

5.2 Navigation Accuracy Validation

To evaluate the precision of the waypoint navigation system, the distance between the UGV's measured arrival coordinates and the target waypoints is computed. This study employs the Haversine formula to determine the

Figure 5. Satellite map view of the navigation experiment at Taipei Port.

great-circle distance between two points on the Earth's surface. To ensure mathematical conciseness and maintain appropriate page margins, the calculation is performed using an intermediate variable a, representing the square of half the chord length between the points:

$$a = \sin^2\left(\frac{\Delta\phi}{2}\right) + \cos\phi_1 \cos\phi_2 \sin^2\left(\frac{\Delta\lambda}{2}\right) \tag{3}$$

Subsequently, the final navigation error distance d is derived as follows:

$$d = 2r \arcsin\left(\sqrt{a}\right) \tag{4}$$

In these equations, r denotes the Earth's mean radius (approximately 6,371 km), while ϕ and λ represent the latitude and longitude in radians, respectively. This metric allows for a precise quantitative assessment of the UGV's positioning accuracy relative to the commanded coordinates within the port environment.

As shown in Figure 5, the validation results demonstrate that the UGV achieved a mean navigation error within 0.032 meters across all nine waypoints. The close proximity between each yellow target point and its corresponding blue arrival position confirms the effectiveness of the coordinate conversion module and the GPS-based navigation system in outdoor port environments.

To assess the repeatability of the navigation system, two additional navigation experiments were conducted at Taipei Port under comparable open-sky conditions. Across all three trials, covering a total of 303 waypoints, the mean waypoint navigation error ranged from 0.012 m to 0.032 m and the average path deviation ranged from 0.057 m to 0.068 m. The consistent sub-centimeter to low-centimeter accuracy across trials confirms the stability of the GPS-based navigation system and coordinate conversion module under the designed operating conditions. Future work will extend validation to environments with partial sky obstruction to further characterize system robustness.

5.3 Anomaly Detection Results

The YOLOv8s model was applied to detect transformer boxes in RGB images, enabling targeted thermal analysis of critical infrastructure. In this experiment, the model accurately detected the transformer boxes, achieving a precision score of 0.875 and a recall score of 0.636. The detected bounding boxes were successfully mapped to the corresponding thermal image regions for temperature extraction. The thermal image used to identify the transformer boxes and indicate the temperatures is shown in Figure 6.

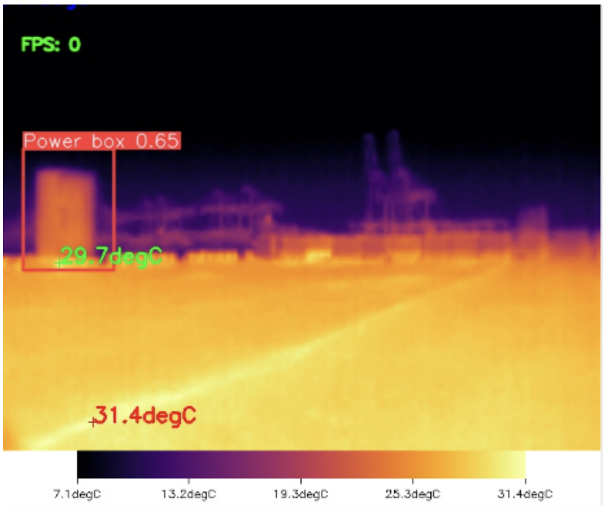

Figure 6. Examples of thermal images for object identification and temperature labeling

The precision of 0.875 confirms that when the model detects a transformer box, the identification is highly reliable, ensuring that thermal analysis is triggered only on correctly localized targets. The recall of 0.636, while moderate, is contextually acceptable: it is consistent with the performance range of YOLO-family models on single-class industrial equipment detection tasks with limited real-world training data, where recall values in the range of 0.60–0.70 are commonly reported due to the scarcity of annotated samples and the visual complexity of real industrial environments. Furthermore, in this application a missed detection results in conservative behavior—the thermal analysis for that asset region is skipped rather than generating a false alarm—making precision the more operationally critical metric. Future work will expand the fine-tuning dataset to improve recall performance.

The gas concentration monitoring module was additionally evaluated during the field experiment. Over the course of the inspection mission, more than ten gas anomaly trigger events were recorded in which the measured concentration exceeded the 25% LEL alarm threshold, and the system successfully detected and transmitted real-time alerts for all recorded events, yielding a detection rate of 100%. This result is consistent with the deterministic nature of threshold-based alarm systems: given a properly calibrated sensor and a well-defined alarm threshold compliant with ISO 10156 and IEC 60079 standards, detection reliability is governed by sensor measurement accuracy rather than classification uncertainty. The event-driven immediate transmission mechanism described in Section 3.3 further ensured that no alarm events were lost due to network latency during the experiment.

5.4 Data Visualization

As shown in Figure 7, the cloud monitoring platform successfully integrated and displayed the multi-sensor data stream in real time. This experiment uploaded nine coordinate points; only two of these points included temperature and gas data, while the rest only provided coordinate information. This is because the UGV only transmits complete temperature and gas data when it detects an anomaly; otherwise, it only transmits coordinates and time data. This allows operators to observe the UGV's inspection path and the location of any anomalies. Red dots on the map interface indicate the locations of the coordinate data transmitted by the UGV. The anomaly alarm message box displays temperature data, combustible gas data, and a site image, allowing operators to observe the data and understand the site conditions. Furthermore, the platform can specify waypoints for the UGV, and the resulting list of coordinates is transmitted to the vehicle via polling-based mechanism described in Section 3. This function enables dynamic task adjustments without physical contact with the vehicle, allowing for flexible adjustments to the inspection path.

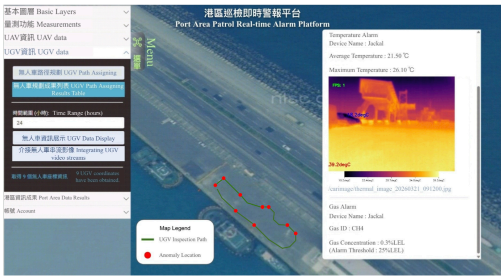

Figure 7. Screenshot of platform interface showing real-time monitoring view

6 Conclusion and Future Work

This research proposes an automated safety inspection system that integrates a UGV with multimodal sensors for autonomous outdoor inspection in port environments. The system addresses key challenges in industrial safety inspection by unifying heterogeneous sensor data through a ROS-based architecture, enabling GPS-based waypoint navigation in unknown environments, and providing operators with an intuitive remote monitoring platform for real-time situational awareness.

Field experiments conducted at Taipei Port validated the system's core performance. Under open-sky conditions, the UGV successfully completed navigation to designated waypoints with a mean waypoint navigation error of 0.032 meters, an average path deviation within 0.068 meters, and GPS positioning accuracy within 0.05 meters. The obstacle avoidance module effectively avoided various obstacles including parked vehicles, electrical boxes, and guardrails without any collisions. The YOLOv8s-based thermal anomaly detection module achieved a precision of 0.875 and recall of 0.636 in identifying transformer boxes, enabling targeted temperature monitoring of critical infrastructure. The gas concentration monitoring module successfully detected all recorded anomaly events (detection rate: 100%), demonstrating reliable threshold-based alarm performance. The cloud-based monitoring platform successfully integrated and displayed multi-sensor data streams in real time, allowing operators to observe inspection progress and anomaly locations.

Despite these results, several limitations require attention. The current system requires manual alignment with true north during initialization due to the lack of an onboard magnetometer, and GPS accuracy degrades under obstructed sky conditions. Navigation repeatability was validated across three trials (303 waypoints total), with mean waypoint navigation error of 0.012–0.032 m and average path deviation of 0.057–0.068 m.

Future work will focus on: (1) integrating a magnetometer or IMU-based heading estimation to eliminate manual initialization; (2) fusing GPS with inertial data for positioning under reduced satellite visibility; (3) expanding the detection module to recognize vehicles, vessels, and storage tanks; (4) adding autonomous return-to-base and battery monitoring; and (5) exploring multi-robot coordination for efficient large-area coverage.

References

[1] Yang Zhang, Yanjun Zhou, Hehua Li, Hao Hao, Weijiong Chen, and Weiwei Zhan. The navigation system of a logistics inspection robot based on multi-sensor fusion in a complex storage environment. *Sensors*, 22(20):7794, 2022. doi:10.3390/s22207794.

[2] David L. Page, Yohan Fougerolle, Andreas F. Koschan, Andrei Gribok, Mongi A. Abidi, David J. Gorsich, and Grant R. Gerhart. SAFER vehicle inspection: A multimodal robotic sensing platform. In *Unmanned Ground Vehicle Technology VI*, volume 5422 of *Proceedings of SPIE*, pages 549–560. SPIE, 2004. doi:10.1117/12.542808.

[3] Qiang Yang, Jingze Dong, Minyao Tan, Jiawei Wang, Dequan Guo, Hao Kang, and Ping Wang. A novel navigation assistant method for substation inspection robot based on multisensory information fusion. *Journal of Advanced Research*, 77:407–418, 2025. doi:10.1016/j.jare.2025.01.016.

[4] Bahador Khaleghi, Alaa Khamis, Fakhreddine O. Karray, and Saiedeh N. Razavi. Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, 14(1):28–44, 2013. doi:10.1016/j.inffus.2011.08.001.

[5] Dongjie Wu, Xunyu Zhong, Xiafu Peng, Huosheng Hu, and Qiang Liu. Multimodal information fusion for high-robustness and low-drift state estimation of UGVs in diverse scenes. *IEEE Transactions on Instrumentation and Measurement*, 71:1–15, 2022. doi:10.1109/TIM.2022.3205687. Art. no. 8505115.

[6] Dongjie Li, Huaitao Shi, Bing Cai, Xiaotian Bai, Lu Chen, Gang Han, and Chenghong Mi. A review of technical advances and applications of intelligent inspection robots in structural health monitoring. *SmartBot*, 1:e70000, 2025. doi:10.1002/smb2.70000.

[7] Shekhar Suman Borah, Aaditya Khanal, and Prabha Sundaravadivel. Emerging technologies for automation in environmental sensing: Review. *Applied Sciences*, 14(8):3531, apr 2024. doi:10.3390/app14083531.

[8] Mofiyinfoluwa Tobi Olowe and Michael Ayomoh. Development of integrated framework for automated construction progress sensing, monitoring and evaluation. In *Proceedings of the 12th International Electronic Conference on Sensors and Applications (ECSA-12)*, volume 118 of *Engineering Proceedings*, page 49. MDPI, nov 2025. doi:10.3390/ECSA-12-26603.

[9] B.L.E.A. Balasuriya, B.A.H. Chathuranga, B.H.M.D. Jayasundara, N.R.A.C. Napagoda, S.P. Kumarawadu, D.P. Chandima, and A.G.B.P. Jayasekara. Outdoor robot navigation using gmapping based slam algorithm. pages 403–408, 2016. doi:10.1109/MERCon.2016.7480175.

[10] S. K. Manoharan, R. K. Megalingam, and A. V. Anilkumar. Autonomous navigation through gmapping-based slam: A comprehensive evaluation. In *2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT)*, 2024. doi:10.1109/ICCPCT61902.2024.10673228.

[11] Javier Sanchez-Cubillo, Javier Del Ser, and José Luis Martin. Toward fully automated inspection of critical assets supported by autonomous mobile robots, vision sensors, and artificial intelligence. *Sensors*, 24(12):3721, jun 2024. doi:10.3390/s24123721.

[12] Hao Jiang, Zijian Guo, and Zhongyi Zhang. Digital twin and path planning for intelligent port inspection robots. *Journal of Marine Science and Engineering*, 14(2), 2026. doi:10.3390/jmse14020186.

[13] Muhayy Ud Din, Waseem Akram, Ahsan B. Bakht, and Irfan Hussain. Llm-vlm fusion framework for autonomous maritime port inspection using a heterogeneous uav-usv system, 2026. URL `https://arxiv.org/abs/2601.13096`.