Qualitative Analysis of RGB-D SLAM Using RTAB-Map for Indoor and Outdoor GNSS-Denied Environments

Ryan Satria Wijaya [1]*, Riski Yosepha Sihombing [2]**
* Department of Electrical Engineering Batam State Polytechnic, Batam, Indonesia
ryan@polibatam.ac.id [1], riskiyosepha@gmail.com [2]

Article Info

Article history:

Received …
Revised …
Accepted …

Keyword:

Loop Closure, GNSS-denied Environments, RGB-D SLAM, RTAB-Map, Visual Inertial Odometry.

ABSTRACT

This study presents a qualitative and semi-quantitative analysis of the Real-Time Appearance-Based Mapping (RTAB-Map) RGB-D SLAM framework utilizing an Intel RealSense D435i sensor running on ROS 2 Humble across both structured indoor and unstructured outdoor GNSS-denied environments. Operating within a controlled 4.5 m x 3.5 m closed-loop path at a fixed sensor height of 0.75 m, the system's performance was evaluated based on visual odometry stability, 3D point cloud reconstruction quality, loop closure efficacy, and computational resource consumption. The experimental results demonstrate that the indoor environment provides highly favorable conditions characterized by rich geometric textures and stable artificial illumination, enabling the Visual-Inertial Odometry (VIO) pipeline to achieve robust tracking, two confirmed returns to the start, and an exceptionally low final drift of 0.025 m with a peak CPU usage of 25.8% and 4.8 GB of RAM. In contrast, the outdoor environment is very challenging, with ambient sunlight causing significant active infrared depth sensor saturation, dynamic vegetation and moving obstacles. This results in unrecoverable tracking losses, 0 successful loop closures, a final drift above 5.00 m, and increased computational overhead reaching 41.8% peak CPU and 6.5 GB of RAM. The results demonstrate the reliability and efficiency of pure RGB-D SLAM for indoor navigation, but also expose the intrinsic limitations of this method in uncontrolled outside daylight environments, thereby hinting the necessity for multi-sensor fusion techniques in the future.

I. Introduction

The capacity of a robot to concurrently sense its surroundings and determine its position is crucial to autonomous navigation [1] . Although GNSS offers global location in outdoor contexts, it is intrinsically unavailable or extremely unreliable in subterranean buildings, thick urban canyons, and indoor environments [2] , as extensively documented across recent visual SLAM literature [3]. As a result, the de facto solution for GNSS-denied navigation is Simultaneous Localization and Mapping (SLAM).

RGB-D cameras provide a compelling combination between direct metric depth perception and rich visual elements across several sensor modalities [4]. A well-known graph-based SLAM technique called RTAB-Map (Real-Time Appearance-Based Mapping) combines appearance-based loop closure detection with visual odometry [5] . Despite its extensive use, there are still few thorough assessments in the literature that describe its performance constraints, particularly when comparing controlled indoor environments with demanding outdoor circumstances employing pure RGB-D inputs [6].

This work aims to bridge this gap by using an Intel RealSense D435i to do a thorough qualitative and semi-quantitative analysis of RTAB-Map. This work makes three main contributions: (1) a comprehensive analysis of visual odometry stability and loop closure efficacy in both indoor and outdoor GNSS-denied scenarios; (2) a study of sensor-specific limitations, particularly the impact of ambient sunlight on active infrared depth sensors; and (3) a semi-quantitative analysis of trajectory drift and computational resource consumption on edge-compatible hardware.

e-ISSN: 2548-6861

II. METHOD

This section describes the overall process of the design, implementation, and evaluation of the proposed RGB-D SLAM system. It starts with the system architecture in general and the hardware configuration used for synchronized data gathering. Then, the integration of the RTAB-Map algorithm, for incremental mapping and loop closing, is detailed. Finally, a controlled experimental setup, semi-quantitative evaluation metrics, and computational resource monitoring methods are described, enabling a rigorous and repeatable assessment of the system's performance in different GNSS-denied scenarios.

A. System Architecture

Figure 1. The system configuration of the proposed RGB-D SLAM architecture

The whole architecture for RGB-D SLAM functions running in the ROS 2 framework [7] is depicted in Figure 1. The system is composed of four main stages, which are data acquisition by the sensors, computation of visual inertial odometry (VIO), SLAM using RTAB-Map and 3D map visualization. The Intel RealSense D435i camera is the RGB-D sensor [8], which perceives the environment in real-time at the hardware level. It generates three types of synchronized data: RGB images, depth images, and camera information, as well as Inertial Measurement Unit (IMU) data. Therefore, this data is published using ROS 2 middleware with topics such as /camera/color/image_raw, /camera/aligned_depth_to_color/image_raw, /camera/color/camera_info and /camera/imu as channels for the nodes to communicate.

The initial step in the process is to fuse the RGB and depth pictures with the high frequency IMU data via the odometry node. This visual-inertial approach allows systems to extract visual cues, back-project 2D features into 3D metric scale coordinates and estimate the related camera pose (rotation and translation) more reliably [9] . The integration of inertial data is important to fill in the small disruptions of visual features and to stabilize fast camera movements. The fast and precise analysis of 3D points in each frame may help the system find the movements between frames and avoid real-time tracking failure. Then the location estimate results of the camera are published on the /odom topic. For incremental mapping, the back-end RTAB-Map SLAM node then subscribes to positional information from the /odom topic. This node also knows when a loop has been completed, by comparing visual fingerprints from previously visited places. It also improves the pose graphs to correct the accumulated front-end drift [10] . Additionally, to maintain responsiveness in long mapping sessions, the system uses an effective memory management strategy to discard unnecessary keyframes. The front-end focusing on fast visual-inertial tracking may be decoupled from the back-end dealing with resource-hungry activities such as pose graph optimization and loop closure detection, and the system can still perform well [11]. Similar to lightweight loop closure schemes proposed for resource-constrained platforms [12] . Our segmented design also guarantees a lightweight and responsive system with accurate 3D point cloud maps and trajectory representations.

B. Hardware Configuration

The proposed framework may collect ambient data only using the Intel RealSense D435i. This apparatus was selected for its capacity to provide hardware-synchronized color and depth streams essential for achieving metric-scale localization without outdoor GNSS signals. The RGB-D camera, Intel RealSense D435i, used in the study is seen in Figure 2. The primary operational characteristics are detailed in Table 1.

Figure 2. The Intel RealSense D435i RGB-D Camera was Used in the experiment

TABLE I
CAMERA SPECIFICATION

Parameter	Value
Camera Type	Intel RealSense D435i
Resolution RGB	1280 x 720
RGB Format	RGB8
Resolution Depth	848 x 480
Depth Format	Z16
Frame Rate	30 FPS
Sensor Type	RGB-D with integrated IMU

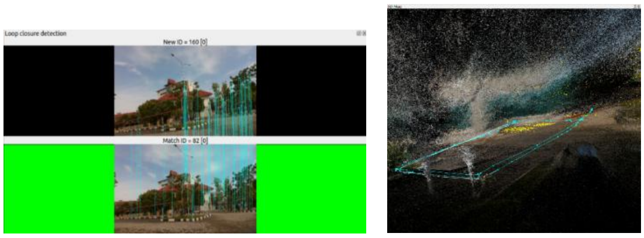
Camera Specification The stereo module was designed to generate depth data in Z16 format with a resolution of 848 × 480 pixels, and the stream from the RGB camera was configured to RGB8 format with a resolution of 1280 × 720 pixels. Both streams were synchronized at the hardware level and captured at a constant frame rate of 30 FPS to provide stable input to the visual odometry process. Both modules were set to auto-exposure to respond to the different lighting conditions seen during indoor and outdoor navigation. The built-in Inertial Measurement Unit (IMU) of the Intel RealSense D435i was actively included in the SLAM process to complement the optical odometry. High-frequency motion estimates from the inertial data provide robustness to rapid camera movements and compensate for the gaps during transitory visual feature losses. Such a visual-inertial fusion approach improves the system's capability to preserve tracking continuity, notably in adverse conditions when visual features might be briefly degraded [13].

C. RTAB-Map Algorithm

The major input of the RTAB-Map architecture is the odometry calculated by the RGB-D visual odometry module. In this formulation, visual odometry serves as a first stage motion estimator, giving the back-end SLAM engine with incremental pose estimations from consecutive sensor observations. For this study, RTAB-Map is chosen as SLAM framework because to its efficient memory management for large-scale mapping applications and its capability to conduct loop closure detection in real-time.

It builds up a graph-like representation of the surroundings, using information from visual odometry, in a sequential manner. A bag of words approach is used to match newly observed visual data with previously stored visual signals for loop closure detection. As can be seen from Figure 3(a), the system is able to identify the previously visited areas by comparing the present visual attributes (New ID = 160) with the recorded keyframes (Match ID = 82). The approach improves the global trajectories consistency and reduces the cumulative drift of incremental visual odometry for long-term mapping [3]. To enhance the computational efficiency of RTAB-Map in extended sessions, a memory management method is employed to maintain a limited working memory (working set of current keyframes) for real-time processing, while moving older keyframes to long-term memory for loop closure detection.

RTAB-Map is a sophisticated backend that operates autonomously to optimize the robot trajectory and give a globally consistent map. It does not let errors accumulate slowly, but considers the whole trip and cancels the whole deviation. This is done by weighting the sensor data differently, giving readings with high confidence a higher weight than those with low certainty. The system easily fuses this updated information with the RGB-D camera inputs to generate a separate and accurate 3D representation of the environment as the trajectory and path change. The resultant three-dimensional map presented in Figure 3(b) depicts the spatial layout of the environment and the effectiveness of the SLAM optimization. The continuity and stability of the estimated motion are evaluated qualitatively by examining trajectories [14]. Discontinuities or anomalies can signal a larger build-up of drift, or less faith in the motion estimates. Odometry estimates are stable and trajectories are smooth and continuous.



(a) (b)
Figure 3. RTAB-Map Visualization

Figure 3 RTAB-Map visualization in outdoor experiments: (a) loop closure detection shows the ability of the system to recognize previously visited locations, as a loop closure is successfully established between New ID = 160 and Match ID = 82; (b) the built 3D point cloud map with the estimated trajectory (cyan line) shows the loop closure adjustment.

D. Experimental Setup

The system's adaptability was tested through experiments in two different GNSS-denied environments: an unstructured outdoor environment along an unpaved pathway adjacent to the Techno Entrepreneur building at the Politeknik Negeri Batam campus and a structured indoor environment in an indoor corridor of the Barelang Robotics Artificial Intelligence Lab (BRAIL) [15] . In each case the robot followed a consistent trajectory of 4.5 x 3.5 meters to provide a controlled and fair evaluation. Data was collected using a portable set-up to simulate the mobile platform moving across the environment. The Intel RealSense D435i was firmly fastened to a vertical framework of 0.75 m to keep the camera height fixed and steady throughout the movement. Moreover, the typical handshake artifacts of informal handheld recording were significantly reduced by clamping the stand tightly to eliminate vertical movement and provide a constant trajectory. All SLAM calculations were performed on a mid-range laptop running Ubuntu 22.04 LTS and ROS 2 Humble, with an Intel Core i7 CPU and 16 GB of RAM. This hardware configuration has been purposefully selected to evaluate the real-time performance of the system on common computing platforms with modest resources, without using expensive GPUs [15].

E. Evaluation Metrics

Quantitative measures like Absolute Trajectory Error (ATE) and Relative Pose Error (RPE) [16] could not be computed owing to the lack of an accurate ground-truth trajectory in the unstructured indoor and outdoor surroundings. Therefore, the system performance is assessed by a hybrid approach, comprising qualitative and semi-quantitative measures, i.e., the success rate of loop closure, the trajectory coherence, the structural density of the reconstructed 3D point cloud, and the final trajectory drift estimated from the odometry log data. This technique thoroughly evaluates the spatial robustness of the proposed RGB-D SLAM system in real-life situations without GNSS.

e-ISSN: 2548-6861

In an effort to provide quantitative context to the qualitative trajectory evaluation, we define trajectory drift as the Euclidean distance between the original starting position and the final estimated location in the X-Y plane, as illustrated below:

$$Drift = \sqrt{(x_{end} - x_{start})^2 + (y_{end} - y_{start})^2} \qquad \text{Equation 1}$$

This equation provides a simple gauge to determine the cumulative positional error during the whole trip. In a closed-loop experiment, the beginning and ending coordinates must be the same. Therefore, any predicted non-zero drift is the uncompensated error integrated by the front-end odometry. A smaller drift value means better odometry accuracy and excellent loop closure corrections, whereas a larger value means a significant degradation of the tracking [17].

The standard deviation of the vertical (Z-axis) position is also examined to establish the vertical stability of the depth estimate. A lower Z-axis standard deviation means that the floor detection was constant, while a higher standard deviation shows that the vertical variations increased owing to the noise from the outdoor depth sensor [18].

F. *Monitoring of Computational Resources*

During all testing runs, we evaluated the computational resource use to confirm the system's practicality for real-time deployment on resource-limited platforms. Essential performance metrics such as CPU usage, RAM use and swap memory consumption were recorded using ROS 2 diagnostic tools and traditional Linux profiling applications, particularly htop [5] . This empirical profile intends to examine the processing overhead of the decoupled front-end and back-end architecture [19] . The computational characteristics are then analyzed in detail and plotted in Section IV (Results and Discussion).

III. Results and Discussion

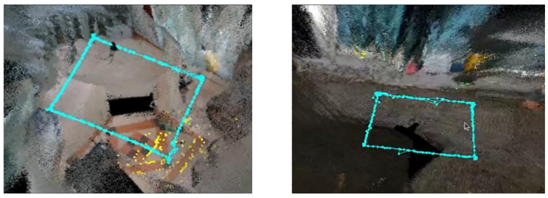
In this part, we provide the results of the assessment of the proposed RGB-D SLAM framework using qualitative and semi-quantitative analyzes. The trials were performed in a field of precisely 4.5 x 3.5 meters, with the RGB-D sensors at a fixed height of 0.75 meters above the ground but in two different configurations, one indoor and one outdoor without GNSS signals, so as to get a balanced and quantifiable result. We controlled for camera distance traveled and elevation to ensure that any differences in the findings were really a function of factors in the environment and not only a function of distance accumulation. The analysis was performed qualitatively, taking into account four main aspects of the RTAB-Map results: visual odometry reliability, 3D mapping accuracy, loop closure, and computing resources.

A. *Visual Odometry Reliability*

(a) (b)

Figure 4. Real-world surroundings including labels A, B, C, D, directional indicators, and measurements of 4.5 m by 3.5 m.

Figure 4 depicts a regulated rectangular closed-loop trajectory (4.5 m x 3.5 m) with sharp corners under actual test conditions for (a) indoor and (b) outdoor scenarios. The trajectory corners (A→B→C→D→A) are marked by points A, B, C and D; for the duration of the data collecting, the RGB-D sensor was positioned at a constant height of 0.75 m from the floor or ground surface, which means that ideally the 3D trajectory should be a straight horizontal line on the z-axis (no up and down movement). This stringent regulation of the sensor height and field dimensions guarantees that variations in indoor and outdoor results are only caused by environmental factors rather than changes in experimental conditions.



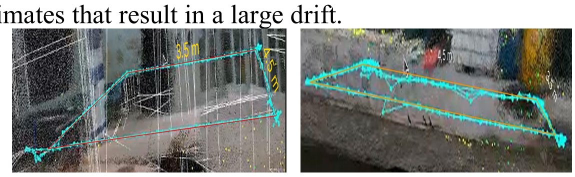
(a) (b)

Figure 5 Top side comparison of trajectories

The top-view assessment results (X-Y axis) are shown in Figure 5. The resulting trajectories have high geometrical precision in the indoor environment (Figure 5a), with sharp corners and rectangular shapes. The key reason for its success is the availability of consistent and rich visual qualities. To extract the feature points, the Visual Inertial Odometry (VIO) approach largely leverages texture and corners. The contrasting geometric patterns on the floor and the sharp angles of the furniture give thousands of static reference points that allow the camera to precisely track movement.

With these powerful qualities, the RGB-D sensor has distinct physical limitations in this situation. In addition to the glare on the floor and the bright TV screen, there are transparent glass dividers with an IR light that causes depth noise and a dark colored sofa that absorbs IR light and reduces the quality of depth readings in this location. This mixture of conditions could lead to local drift and missing areas of the point cloud, especially when moving things such as people walking by are present. However, the overwhelming dominance of static characteristics (textures and corners) makes RTAB-Map able to neglect local noise caused by reflections, transparency and infrared absorption. Loop closure and memory management approaches are able to correct for these interruptions well, leading to a minimal level of total drift accumulation.

However, in the outdoor setting (Figure 5b), a considerable decrease in performance is seen. The reconstructed trajectory has a clear horizontal drift, and is in the shape of a curved rectangle with rounded corners. Some of the harsh environmental conditions that lead to this inaccuracy include the existence of naturally dynamic objects such as trees and plants moving in the wind, loss of continuous visual characteristics owing to sunlight interferences and the erratic movement of autos and humans. Both the vegetation motion and these transient human and vehicle activities violate the static world assumption of visual-inertial odometry. Due to the noise supplied by these dynamic elements, the VIO algorithm obtains biased motion estimates that result in a large drift.



(a) (b)

Figure 6. Comparison of trajectories, side view (3D elevation / Z-axis)

Figure 6 illustrates a side-by-side analysis of paths along the Z axis. (a) Indoor: Consistent height assessment with little geographical variations. Outdoor: Large vertical shift due to sun radiation and deterioration of depth. Figure 6 might yield considerable insights when examined from a lateral viewpoint (Z-axis). The optimal route is a straight horizontal line, if the camera moves across a level surface at a consistent height of 0.75 m. Within the indoor environment, the trajectory is quite uniform (Fig. 6a). The method avoids sudden variations or spikes, even while moving across flat, unadorned terrains. This may be achieved when the camera viewpoint regularly includes more reliable reference points such as the irregular floor and prominent furniture edges, guaranteeing stable tracking and a fluid trajectory. Slight modifications occur, mostly in the BRAIL corridor, where the depth sensor is temporarily misled by people walking by or reflections from the ground.

Conversely, the outer shell (Figure 6b) exhibits significant unpredictable variations in both ascending and descending trajectories. The path is not linear. It contorts and fluctuates in several unexpected manners. The persistent variation in changing light from intense sunshine to total darkness leads to considerable vertical and pitch inconsistencies, along with unpredictable infrared washout patterns. The rapid changes in lighting and the lack of stable visual components lead to significantly erroneous depth information for the RGB-D sensor.The system cannot tell the difference between actual movement and sensor disturbance and it reads these changes as either a camera movement or a tilt.

B. Qualitative Analysis of the Reconstruction 3D Point Clouds

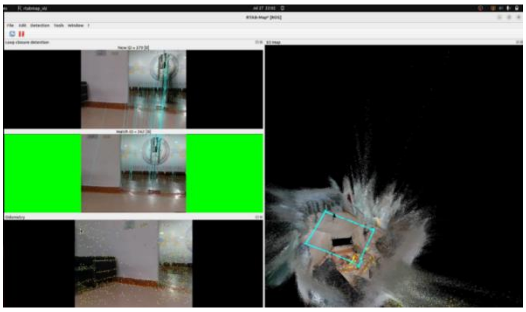


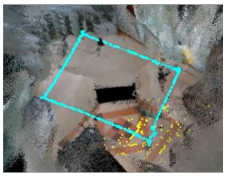
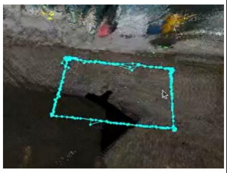
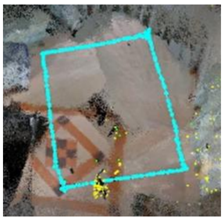
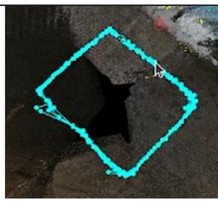
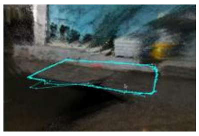
Figure 7. A reliable projected path (cyan line) and comprehensive structural mapping are presented in an indoor 3D point cloud reconstruction.

Fig 7. Indoor 3D point cloud reconstruction of the BRAIL laboratory corridor. The indoor mapping displays a great structural fidelity, as well as the generation of a dense point cloud. The RTAB-Map could provide a consistent 3D representation of the corridor with consistent depth measures and dependable VIO. Doors, walls, furniture etc. are all different geometrical limitations. The loop closure approach was quite successful at matching overlapping keyframes, evidenced by the superb consistency of the texture mapping and the absence of “ghosting” or double-edge aberrations. This demonstrates the efficiency of the memory management and keyframe selection in structured indoor environments.

Figure 8. Present a fractured course (cyan line) and sparse, spatially disordered structure mapping.

In contrast, the outdoor 3D reconstruction (Figure 8) is significantly fragmented, less abundant, and spatially disordered. The degradation is directly attributed to the odometry instability and sensor limitations discussed earlier. The generated map reveals a sequence of cascading failures as the system translates 2D RGB pixels into 3D space thru the use of noisy depth maps and compromised pose estimates, which are worse by substantial Z-axis changes and sporadic tracking disruptions. These complications appear as substantial unmapped voids, floating dots, and "ghosting" phenomena where ambient sunlight has saturated the active infrared pattern. Moreover, the mapping back-end cannot integrate observations into a unified global map because of dynamic elements such as moving shadows and oscillating vegetation, leading to inconsistent feature correspondences. In conclusion, this highlights the difficulties encountered by pure RGB-D SLAM in producing a metrically accurate 3D model under bright, chaotic outdoor conditions.

TABLE II
QUALITATIVE ANALYSIS OF SLAM TRAJECTORIES FROM INDOOR AND OUTDOOR TRIALS

Environment	Top view	Side view	Shape Field
Indoor (1)			Rectangular form with sharp corners. Feature tracking confidence is good, Z axis drift is low.
Outdoor (1)			Moderate geometric distortion. Rounded edges. Big changes in the Z axis are due to sunlight interference.
Indoor (2)			High repeatability with Trial 1. Stable elevation with rigorous adherence to the 4.5 m × 3.5 m closed-loop route.
Outdoor (2)			Consistent deterioration trend with Trial 1. Confirms systemic limits of the active IR sensor in non-structured lighting.

The tests were performed for each setting in two trials (Trial 1 and Trial 2) to examine the consistency and repeatability of the proposed framework. The visual comparison of the predicted trajectories directly retrieved from the RTAB-Map interface is provided in Table 2. In the indoor scenarios (Rows 1 and 3), both trials show significant geometric purity, with a sharp rectangular shape centered precisely at the specified 4.5 m $\times$ 3.5 m path. The side views reveal a low drift in the Z-axis, indicating that the Visual-Inertial Odometry (VIO) pipeline is tightly locked to the rich static features of the structured environment. In contrast, the outdoor circumstances (Rows 2 and 4) exhibit large geometric aberrations in both trials. The top views show route deviations and rounded corners. The side views exhibit jagged oscillations in the Z-axis. The very similar degradation trend of both outdoor trials proves that the trajectory errors are not random outliers but a systematic impact of the saturation of infrared sensors and the dynamic limitations of the environment.

C. Loop Closure Evaluation

To be able to constrain the accumulated drift of the visual odometry, loop closure detection is required. In this work, RTAB-Map uses the bag-of-words approach to match recorded keyframes with the current visual signatures. Once a successful match is made pose graph optimization is performed, to make our assessment metrics clear, it is vital to be able to distinguish between a general loop closure and a return to start. A "Return to Start" requires the system to identify the precise start position (Node 0) after finishing a lap. A general loop closure can happen at any previously visited place (e.g. half-way along the trajectory). The operator followed the same 2-lap rectangular predetermined course in both settings of our closed-loop experimental design. The camera was physically panned around the starting area 3 times, once at the start, once at the end of Lap 1 and once at the end of Lap 2. Hence to avoid confusion we characterize this measure properly as “Confirmed Loop Closures at Start Position”. We only measure the ability of the algorithm to match the current visual signature to the beginning keyframe at the specified start point to ensure that this metric measures the score’s ability to correctly re-localize globally and not simply to large, arbitrary loop closures mid-path or the operator’s choice of path. A detailed comparison of the experimental outcomes is shown in Table 3.

TABLE III
CONTRAST BETWEEN FINDINGS FROM INDOOR AND OUTDOOR TRIALS

Parameters	Indoor	Outdoor
Duration (min)	14.22	13.58
Odometry frames	2,890 frames	1,940 frames
Traversal Laps	2	2
Outlier Events (Cov = 9999.0)	2 (Transient/Isolated)	Multiple (Continuous Blocks)
System Recovery Capability	High (Successfully recovered)	None (Unrecoverable tracking loss)
Lap 1 – Max Drift (m)	4.40	> 6.00
Lap 1 – End Drift (m)	0.04	3.50-4.00
Lap 2 – Max Drift (m)	4.40	6.50
Lap 2 – End Drift (m)	0.025	5.82
Confirmed Returns to Start	2	0
Overall Final Drift (m)	0.025 (after optimation)	> 5.00
Z-axis Std Dev (m)	0.18 very stablish	> 0.30
Feature Tracking Stability	High	Low / Unstable
Loop Closure Performance	Highly Successful (2 Returns)	Unsuccessful

The main qualitative problems for the failure of outdoor loop closure are the large variation in lighting and the lack of distinctive and stable visual cues in unstructured environments, which can seriously impair the robustness of visual feature matching. Consequently, the outdoor trajectory experienced an unrectified deviation throughout the excursion. In conclusion, this comparative research reveals that outdoor conditions provide complex challenges that together undermine odometry stability and re-localization skills, while indoor settings offer highly favorable conditions for RGB-D SLAM loop closure. The numerical examination of the indoor odometry records reveals a significant stability of the system. The VIO algorithm gave very accurate pose estimates across the whole 14.22 min, which consists of 2,890 odometry frames with covariance measures continuously < 0.001 m^2. The loop closures were identified correctly as the end of each traversal lap, yielding two confirmed returns to the initial location. The back-end pose graph optimization, thus, effectively fixed the little differences in the front end and the overall drift was very small, about 0.025 m. Furthermore, the vertical variation of the camera elevation was small, as the standard deviation along the Z-axis was tightly constrained to roughly 0.18 m.

Forensic analysis of the raw odometry data reveals a clear degradation in tracking quality. A covariance value of 9999.0 is the universal indicator indicating erroneous or indeterminate data (full tracking failure) in ROS 2. Instances of 9999.0 covariance in the indoor case were short and few, typically only happening during fast camera motion or sensor initialization. Once the visual characteristics were stabilized, the VIO pipeline showed a great level of robustness, recovering rapidly and obtaining extremely reliable covariance values (e.g. < 0.001 m2). For the outdoor environment, there were long periods with 9999.0 covariance values and null position estimates (x:0.0 y:0.0 z:0.0). Therefore, the surrounding sunshine strongly occluded the active infrared sensor, and the VIO module lost spatial fixation for long periods of time. In the latest 3D reconstruction, a large drift in the trajectory was seen. This was a direct outcome of the system's inability to recover from the outdoor tracking failures (the opposite of the indoor case).

All raw odometry data were recorded in real-time during the physical trials using the ROS 2 rosbag2 architecture to assure empirical validity and reproducibility. The ongoing telemetry information from the VIO node was disseminated on the /odom topic (message type: nav_msgs/msg/Odometry) and was consistently recorded into a SQLite-based .db3 database file along with a corresponding metadata.yaml configuration file. This unprocessed binary log was interpreted and converted into a structured text format for forensic examination, preserving the original Unix timestamps, three-dimensional positional coordinates (X, Y, Z), and the 6x6 pose covariance matrices. Be aware that the acquired data represents the unrefined front-end odometry output that was recorded in real-time from the sensor. Thru the assessment of this unprocessed data, we offer unequivocal mathematical validation of the real-time tracking reliability and intrinsic drift properties of the system, together with assurances that all reported metrics are grounded solely in actual hardware performance, rather than theoretical conjectures.

e-ISSN: 2548-6861

D. Loop Closure Evaluation

Table 4 concludes the previous sections and shows a detailed comparison between the results of the indoor and outdoor experiments. This synthesis explains the large performance gap by focusing on the function of the algorithmic resilience, the physical restrictions of the sensors and the environmental variables.

TABLE IV
DEPTH ASSESSMENT OF SLAM CAPABILITIES FOR INDOOR AND OUTDOOR SETTINGS

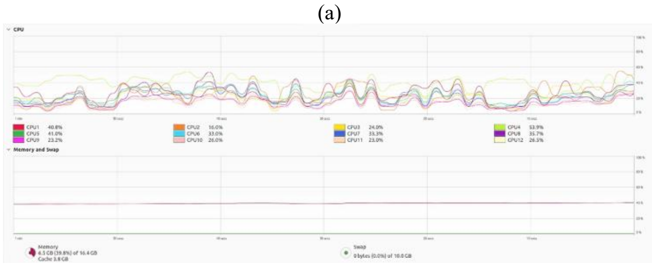
Evaluation Aspect	**Indoor Environment**	**Outdoor Environment**
Illumination Conditions	Stable, artificial lighting; minimal IR saturation.	Strong ambient sunlight; severe IR depth sensor saturation.
Visual Features	Rich and structured (walls, corners, furniture)	Repetitive textures, smooth surfaces, dynamic vegetation.
Pose Covariance	Consistently low (< 0.001 m²).	High / Invalid (spikes to 9999.0 during tracking loss).
Z-axis Stability	Highly stable (Std Dev = 0.18 m).	Severe fluctuations (artifacts dropping to -0.35 m).
Loop Closure Efficacy	Partially successful (2 confirmed returns).	Unsuccessful (0 confirmed returns).
Overall Final Drift (m)	0.025 m post-optimization)	> 5.00 m (unrecoverable drift).
Computational Resources	CPU 25.8%, RAM 4.8 GB (29.5%)	CPU peak 41.8%, RAM 6.4 GB (39.1%).
3D Point Cloud Quality	Dense, structurally coherent, no ghosting.	Sparse, fragmented, significant ghosting artifacts.

The best conditions for RGB-D SLAM are provided by the indoor environment, with its structured geometry, homogeneous artificial illumination and distinguishing high-frequency visual features as seen in Table 4. This enables the Visual-Inertial Odometry (VIO) pipeline to achieve great tracking robustness with low pose covariance. Conversely, the indoor environment is optimal for RGB-D SLAM, as demonstrated in Table 4, due to its organized geometry, homogeneous artificial illumination and the presence of characteristic high frequency visual features. This enables the VIO process to have great tracking robustness and low pose covariance. But the outdoor world is a considerably harder computational and physical task. The active infrared projector of the D435i sensor quickly saturates in ambient daylight, leading to substantial metric-scale aberrations such as the Z-axis position inaccurately falling to -0.35 m. Moreover, the pose covariance spiking to 9999.0 occasionally indicates catastrophic feature tracking losses. The fundamental reasons of these failures are the lack of differentiating textures on outdoor pavements, along with the dynamic environmental conditions such as changing shadows and flora. Outdoor constraints can be mitigated by employing additional sensor modalities (such as LiDAR or advanced multi-sensor fusion), but doing so adds to the computing load. Thus, enhancing the computational economy and environmental robustness of pure RGB-D SLAM frameworks remains a key challenge for further investigation for deployment on resource-limited mobile robotic platforms.

E. Computation Performance

We monitored the computational performance to be able to use the proposed framework in real time on hardware with limited resources (Intel Core i7, 16GB RAM, no dedicated GPU).

(a)



(b)

Figure 9. Computational efficiency during system functionality.

To analyze the computational requirements of our RGB-D SLAM system, we monitor real-time resource utilization on a 12-core workstation with 16.4 GB RAM, as shown in Figure 9. The indoor experiment shows good performance with 25.8% peak CPU and 4.8 GB RAM used (29.5%) (Figure 9a). The system usured stable real-time performance by seamlessly processing 2,890 odometry frames (~3.4 Hz) without relying on swap memory. In contrast, the outdoor session used greater resources (Figure 9b). CPU usage of 41.8% peak CPU and 6.5 GB used RAM (39.8%) were used significantly. Although the system processed a smaller amount of frames (1,940 frames, ~2.4 Hz), it needed to work harder to deal with fluctuating lighting, noisy depth maps and recurring tracking failures. In general, the framework is quite lightweight for indoor use, but in outdoor contexts the demands on the computational resources increase by roughly 30%. The swap graphs in Figure 9 reveal that in both cases the swap memory is not used. Therefore, the system may work in real-time without large memory limitations.

IV. CONCLUSION

In this study, a comprehensive qualitative and semi-quantitative evaluation of the RTAB-Map RGB-D SLAM architecture was carried out in various GNSS-denied situations. We demonstrate the significant effect of outdoor conditions on system performance and computational overhead, using raw Visual-Inertial Odometry (VIO) telemetry and forensic analysis of pose covariance matrices.

The technology exhibited outstanding stability and robustness in the regulated indoor setting. The VIO pipeline successfully made two confirmed loop closures at the beginning location but maintained highly confident pose predictions. This allowed us correct the front-end errors in the back-end pose graph optimization and to achieve an extremely low final drift of only 0.025 m. Moreover, the processing footprint was very efficient, using only 25.8% peak CPU and 4.8 GB of RAM (29.5%), demonstrating its applicability to edge devices with limited resources.

In contrast, the uncontrolled outdoor environment exposed the fundamental physical limitations of active infrared depth sensors. Severe IR washout from ambient sunlight resulted in zero successful loop closures, catastrophic ultimate drift of over 5.00 m, and protracted tracking losses as evidenced by repeated 9999.0 covariance spikes. This degradation also translated into a high rise of the computational cost, since the system was unable to deal with irregular feature matches and noisy depth maps, peaking at 41.8% of CPU and 6.5 GB of RAM utilization (39.8%).

Finally, our work reveals that the pure RGB-D SLAM is a very reliable and computationally effective solution for indoor navigation, although it is intrinsically inadequate for light unstructured outdoor situations. Future works should be directed on fusing of complementing sensor modalities such as LiDAR or advanced multi-spectral cameras and the development of adaptive sensor fusion algorithms to ensure robust autonomous operation in diverse real-world scenarios.

REFERENCES

[1] G. Grisetti, R. Kümmerle, C. Stachniss, and W. Burgard, "A Tutorial on Graph-Based SLAM."

C. Cadena *et al.*, "Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309–1332, Dec. 2016, doi: 10.1109/TRO.2016.2624754.

Y. Dai, J. Wu, and D. Wang, "A Review of Common Techniques for Visual Simultaneous Localization and Mapping," *Journal of Robotics*, vol. 2023, pp. 1–21, Feb. 2023, doi: 10.1155/2023/8872822.

R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, Oct. 2017, doi: 10.1109/TRO.2017.2705103.

M. Labbé and F. Michaud, "RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation," *J. Field Robot.*, vol. 36, no. 2, pp. 416–446, Mar. 2019, doi: 10.1002/rob.21831.

S. Zhang, L. Zheng, and W. Tao, "Survey and Evaluation of RGB-D SLAM," *IEEE Access*, vol. 9, pp. 21367–21387, 2021, doi: 10.1109/ACCESS.2021.3053188.

S. Arshad and G.-W. Kim, "Role of Deep Learning in Loop Closure Detection for Visual and Lidar SLAM: A Survey," *Sensors*, vol. 21, no. 4, p. 1243, Feb. 2021, doi: 10.3390/s21041243.

A. Macario Barros, M. Michel, Y. Moline, G. Corre, and F. Carrel, "A Comprehensive Survey of Visual SLAM Algorithms," *Robotics*, vol. 11, no. 1, p. 24, Feb. 2022, doi: 10.3390/robotics11010024.

A. Tourani, H. Bavle, J. L. Sanchez-Lopez, and H. Voos, "Visual SLAM: What Are the Current Trends and What to Expect?," *Sensors*, vol. 22, no. 23, p. 9297, Nov. 2022, doi: 10.3390/s22239297.

W. Chen *et al.*, "An Overview on Visual SLAM: From Tradition to Semantic," *Remote Sens. (Basel).*, vol. 14, no. 13, p. 3010, Jun. 2022, doi: 10.3390/rs14133010.

H. Bavle, J. L. Sanchez-Lopez, C. Cimarelli, A. Tourani, and H. Voos, "From SLAM to Situational Awareness: Challenges and Survey," *Sensors*, vol. 23, no. 10, p. 4849, May 2023, doi: 10.3390/s23104849.

Y. Shi, R. Li, Y. Shi, and S. Liang, "A Robust and Lightweight Loop Closure Detection Approach for Challenging Environments," *Drones*, vol. 8, no. 7, p. 322, Jul. 2024, doi: 10.3390/drones8070322.

G. A. Acosta-Amaya, J. M. Cadavid-Jimenez, and J. A. Jimenez-Builes, "Three-Dimensional Location and Mapping Analysis in Mobile Robotics Based on Visual SLAM Methods," *Journal of Robotics*, vol. 2023, pp. 1–15, Jun. 2023, doi: 10.1155/2023/6630038.

H. Gao *et al.*, "Enhancing the Localization Accuracy of UAV Images under GNSS Denial Conditions," *Sensors*, vol. 23, no. 24, p. 9751, Dec. 2023, doi: 10.3390/s23249751.

A. R. Sahili *et al.*, "A Survey of Visual SLAM Methods," *IEEE Access*, vol. 11, pp. 139643–139677, 2023, doi: 10.1109/ACCESS.2023.3341489.

S. Song, F. Yu, X. Jiang, J. Zhu, W. Cheng, and X. Fang, "Loop closure detection of visual SLAM based on variational autoencoder," *Front. Neurorobot.*, vol. 17, Jan. 2024, doi: 10.3389/fnbot.2023.1301785.

Z. Xiang *et al.*, "Accurate localization of indoor high similarity scenes using visual slam combined with loop closure detection algorithm," *PLoS One*, vol. 19, no. 12, p. e0312358, Dec. 2024, doi: 10.1371/journal.pone.0312358.

J. Zhu, H. Li, and T. Zhang, "Camera, LiDAR, and IMU Based Multi-Sensor Fusion SLAM: A Survey," *Tsinghua Sci. Technol.*, vol. 29, no. 2, pp. 415–429, Apr. 2024, doi: 10.26599/TST.2023.9010010.

I. Jarraya *et al.*, "Gnss-denied unmanned aerial vehicle navigation: analyzing computational complexity, sensor fusion, and localization methodologies," *Satellite Navigation*, vol. 6, no. 1, p. 9, Dec. 2025, doi: 10.1186/s43020-025-00162-z.