Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666

Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008

JOURNAL OF INFORMATION SYSTEM AND TECHNOLOGY MANAGEMENT (JISTM)
www.jistm.com

REAL-TIME 3D MAPPING AND LOCALIZATION OF PALM OIL TREE FOR HARVEST DATA MANAGEMENT SYSTEM

Marizuana Mat Daud [1*], Zulaikha Kadim[2], Fazlina Mohd Ali[3], Jun Kit Chaw[4]

[1] Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Malaysia
Email: marizuana.daud@ukm.edu.my
[2] MIMOS Berhad, Technology Park Malaysia, Malaysia
Email: zulaikha.kadim@mimos.my
[3] Faculty of Information Science & Technology, Universiti Kebangsaan Malaysia, Malaysia
Email: fazlina.mohdali@ukm.edu.my
[4] Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Malaysia
Email: chawjk@ukm.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 18.09.2023
Revised date: 15.10.2023
Accepted date: 30.10.2023
Published date: 07.12.2023

To cite this document:

Mat Daud, M., Kadim, Z., Ali, F. M., & Jun, K. C. (2023). Real-Time 3D Mapping And Localization Of Palm Oil Tree For Harvest Data Management System. *Journal of Information System and Technology Management*, *8* (33), 91-101.

DOI: 10.35631/JISTM.833008

Abstract:

The palm oil industry, particularly vital in Asian nations such as Malaysia and Indonesia, plays a pivotal role in their economies. Its efficiency hinges on effective palm oil harvest management. This industry is poised for substantial growth, with global palm oil demand projected to surge from 51 million tons to an estimated 120 to 156 million tons in the next three decades. However, it faces significant challenges, notably its heavy reliance on foreign labor, particularly for the labor-intensive tasks of harvesting and assessing fruit bunch maturity, a situation further exacerbated by recent labor shortages due to the COVID-19 pandemic. Therefore, this study proposed a comprehensive approach to 3D mapping and tagging, centered around the detection and localization of palm oil trees. The key methodology involves employing object detection technology to extract the coordinates of these trees. Subsequently, the RTAB-Map is harnessed to precisely localize these identified palm oil trees within a 3D space. The results of this research demonstrate a robust and efficient system for palm oil tree detection and 3D localization. By utilizing object detection and RTAB-Map, we achieve high accuracy in tree identification and localization, laying the foundation for improved management of palm oil plantations. This technology not only facilitates the monitoring and assessment of palm oil tree health but also enhances overall plantation management, resource allocation, and sustainability efforts. The findings herein signify a significant step towards optimizing palm oil cultivation practices and, consequently, fostering environmental conservation and sustainable agricultural practices within the industry.

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008

Keywords:

Palm Oil Tree, Harvest Management, RTAB-map, 3D Mapping, Tree Detection, Tree Localization

Introduction

Harvesting stands out as the most labor-intensive aspect of crop management, encompassing various stages, including scouting for ripe fresh fruit bunches (FFB), cutting, collecting, and gathering them in the field (Ruiz et al., 2017), (Mosquera et al., 2017). The timing of FFB cutting is critical to optimizing productivity since the oil content of the fruits depends on their ripeness or maturity level. Traditionally, this timing has been determined by assessing the number of loose fruitlets (LF) on the ground. For instance, when 1-10 fruits are found on the ground, it signals that the bunch is ripe and ready for harvesting. This task is typically carried out by a dedicated harvest team or individual workers. However, when entrusted to individual workers, the assessment of FFB maturity often relies on their experience, which can diminish over time, leading to reduced consistency and labor productivity (Castiblanco et al., 2010).

By the year 2050, the global population is expected to reach a staggering 9.7 billion, creating a substantial demand for palm oil, estimated to exceed 240 million tons to meet this burgeoning need (Junos et al., 2021). The palm oil production landscape is dominated by Indonesia and Malaysia, together accounting for an imposing 84% of the world's output. In 2018, Malaysia emerged as the second-largest global producer, contributing a significant 27% share, equivalent to nearly 20 million tons annually (Hannah et al., 2022), as depicted in Figure 1. Furthermore, Malaysia holds the distinction of being the world's second-largest palm oil exporter. In 2020, it exported a substantial total of 17,368,865 metric tons of palm products to key markets including China, India, the Netherlands, Turkey, and the USA (MPOC, 2020).

While Malaysia's oil palm agriculture covers an extensive 5.87 million hectares, a marginal 0.6% reduction was observed compared to the preceding year in 2019 (Parveez et al., 2020). This dip can be largely attributed to the disruptions caused by the COVID-19 pandemic, primarily associated with the implementation of movement control orders (MCO) which significantly impeded operations. Furthermore, the industry grappled with a shortage of labor, further retarding palm oil crop production. Consequently, the agricultural sector is undergoing a profound transformation, as it increasingly adopts artificial intelligence (AI) technology. This marks a pivotal shift, as farmers progressively integrate advanced machinery and tools into their operations, enhancing productivity while concurrently reducing production costs.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008

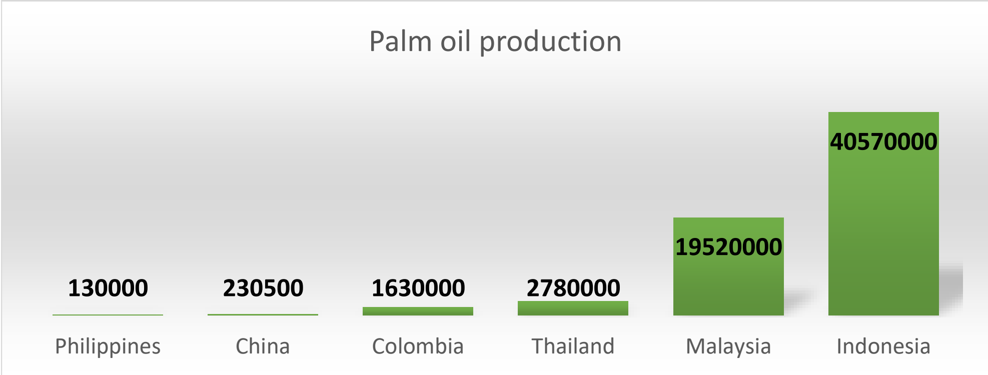


Figure 1: ASEAN Palm Oil in 2018

Lee et al. (2014) introduced a predictive model that specifically named two significant variables, namely the type of smallholder management and the harvesting rotation. Smallholder oil palm yields and incomes were found to be constrained by these variables, respectively. The emphasis on agricultural extension initiatives aiming at distributing optimal management techniques among independent smallholders was one of the main goals achieved through this study project. The study also sought to improve independent smallholders' access to oil palm mills to lower the marketing expenses related to fresh fruit bunches. Considering its consequences for creating a low-income society in Indonesia, the study's findings highlight the crucial function that harvest management plays within the oil palm environment.

Another study by (Zheng et al., 2021), emphasized the importance of implementing an approach incorporating status surveillance and intelligent management of oil palm plantations. This study's main goals were to improve plantation planning, maximize oil palm productivity, and simultaneously lessen the need for physical labor and fertilizer use. Zheng created a novel technique known as the Multi-class Oil Palm Detection approach (MOPAD) to accomplish these goals. This cutting-edge method aimed to provide accurate oil palm tree detection and efficient monitoring of their growth condition. These discoveries emphasize the value of harvest management research on a worldwide scale and its ongoing advancement.

In the realm of visual localization for robotics, the prevalent technologies employed in harvest management primarily encompass Visual Simultaneous Localization and Mapping (VSLAM) and Visual Odometry (VO). A recent investigation conducted by (Diao, 2023) entailed a comprehensive review of the advancements in vision-based navigation and guidance, specifically in the context of autonomous vehicles and robots within agriculture. The study's findings revealed that VSLAM technology predominantly finds application in the realms of visual localization and navigation for agricultural robots.

Consequently, researchers (Sofiah et al., 2023) have created navigation techniques based on vision and lidar that are especially suited for locations without GPS, such as Simultaneous Localization and Mapping (SLAM). This review paper explores the state-of-the-art in visual and lidar-based outdoor navigation without the use of a GPS sensor. To replicate circumstances without GPS coverage, experiments are conducted outside. The study showed that while both LIO-SAM and RTAB-Map use Lidar scans to produce 3D maps, their methods for loop closure detection are different, and this has ultimately helped to improve the monitoring of the outdoor

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008
environment. However, 3D detection using depth camera offers a cost-effective solution with a straightforward setup for object localization in three-dimensional space when compared to LiDAR-based detection systems (Hu et al., 2023).

In light of contemporary research, it becomes apparent that there exist several areas warranting enhancement for the refinement of outdoor monitoring and management, employing diverse Artificial Intelligence (AI) technologies including sensors, SLAM, RTAB-Map, and numerous others. Hence, the primary objective of this study is to delve into real-time 3D mapping and localization rooted in the realm of palm oil tree detection data. Building upon prior research endeavors, which encompassed tree detection (Daud et al., 2023) through deep learning techniques and fruitlet detection (Daud et al., 2022) via a fusion of image processing and a blend of deep learning methodologies, the outcomes of these efforts serve as valuable contributions to the present work. Consequently, plantation managers can utilize this information to efficiently allocate resources for the harvest task, allowing workers to proceed directly to identified trees for bunch cutting. This streamlined approach indirectly contributes to enhanced labor productivity within the plantation.

Methodology

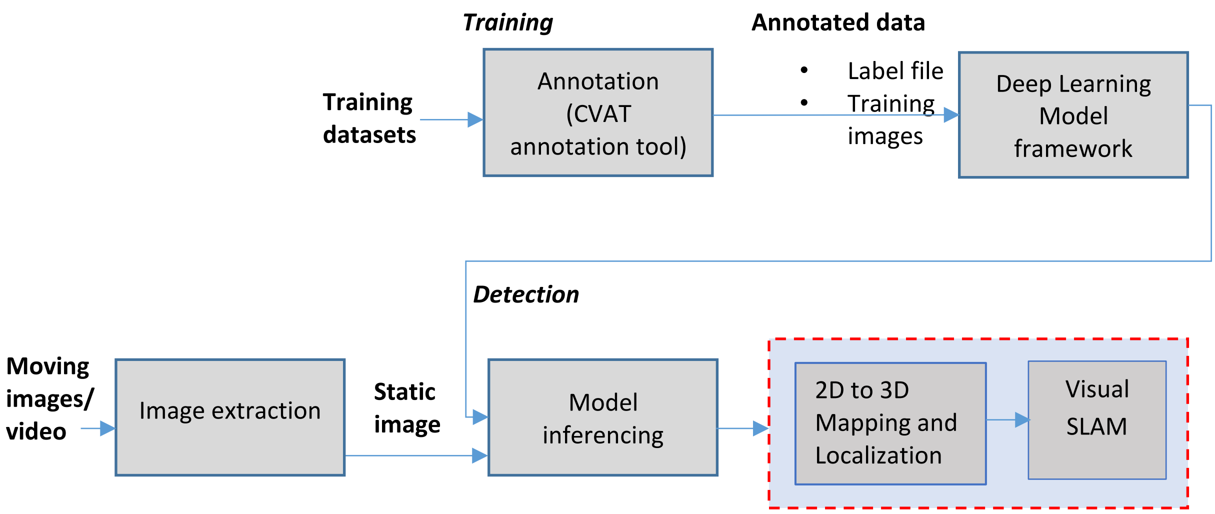
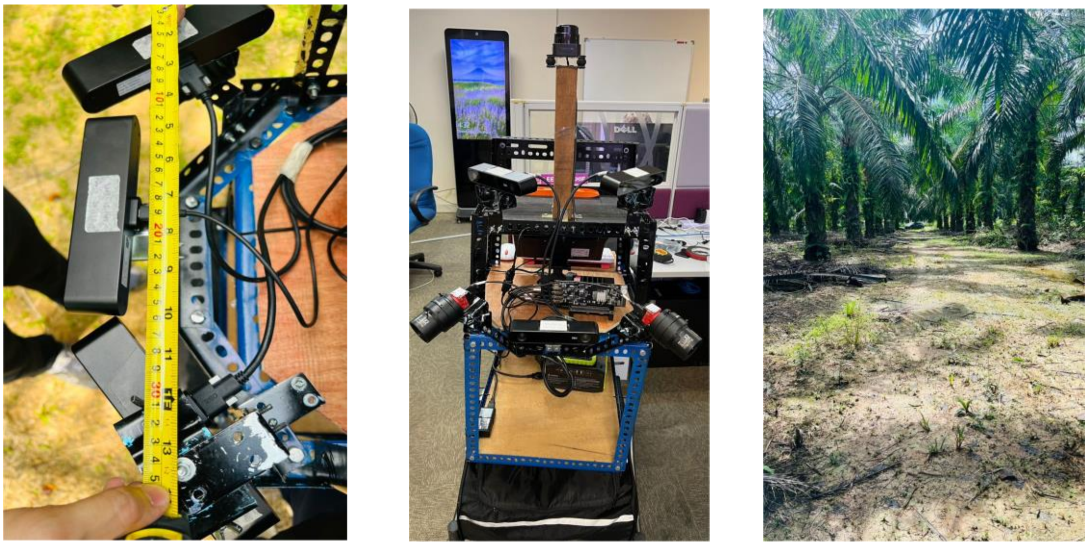


Figure 2: Flow Diagram Of Palm Oil

The system initiation involved the crucial steps of data acquisition, training and testing datasets. The data acquisition part is explained on the next sub-topic. Prior to training, images underwent annotation and labeling using the CVAT annotation tool. In our specific context, the harvest system's functionality hinges on the detection of fruitlets. When the count of fruitlets surpasses a predefined threshold (e.g., 3 fruitlets), the tree is classified as "harvest-ready," and vice versa. Building on our prior research, fruitlet and tree detection were carried out using YOLOv4 and YOLOv5 algorithms. The deep learning model was fed with labeled image files, followed by a series of optimizations. Subsequently, the model was deployed for inference using real-world images. From the detection output, the bounding box's location and coordinates of the tree were extracted, facilitating the transition from a 2D to a 3D representation. This conversion was then mapped onto the 3D mapping using visual SLAM technique, effectively depicting the tree's spatial location. Figure 2 illustrates the workflow encompassing fruitlet and tree detection and localization, with the dotted box delineating the focus of discussion in this paper.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008

Acquiring Image and Camera Calibration



(a) ZED cameras placement

(b) Prototype of Unmanned Ground Vehicle (UGV)

(c) Plantation field

Figure 3: Experimental Setup for Actual Testing in Palm Oil Plantation Area

To determine the precise 3D position of an object, our approach employs a trio of cameras: three ZED cameras positioned strategically to capture distinct viewpoints—left, center, and right, as depicted in Figure 3 (a). The prototype consists of NVIDIA Jetson AGX Xavier series, three ZED cameras, two Alvium camera, battery, monitor and 2D lidar, as shown in Figure 3(b). Left and right ZED camera are used to detect tree on the left and right side, while centre ZED camera is used to do the mapping. The other two cameras are monocular camera (Alvium), primarily used for fruitlet detection. The ZED cameras, also known as binocular or stereo cameras, are equipped with advanced neural stereo depth-sensing technology, capable of delivering depth frames at a remarkable speed of up to 100 frames per second (FPS). This arrangement effectively simulates the left and right perspectives, mimicking human binocular vision. Our chosen object for image acquisition is the palm oil tree (Figure 3(c)), specifically targeted to determine its harvest-ready status. The image capture process involves systematically traversing the plantation, following each row in a loop. A consistent distance of approximately 2 to 3 meters is maintained between the prototype and the tree.

It is crucial to initiate the data collection process by allowing a brief pause of approximately 5 seconds before embarking on the walk. This interval allows for the synchronization and activation of all cameras, ensuring comprehensive environmental capture. During the data collection walk, it is imperative to maintain a steady and controlled pace—not excessively fast, nor excessively slow—while also making a concerted effort to sustain both speed and a centered position throughout the journey. In instances where ground-level obstacles are encountered, the approach involves minimizing trolley elevation, keeping it as close to the ground as feasible to facilitate unhindered progress.

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008

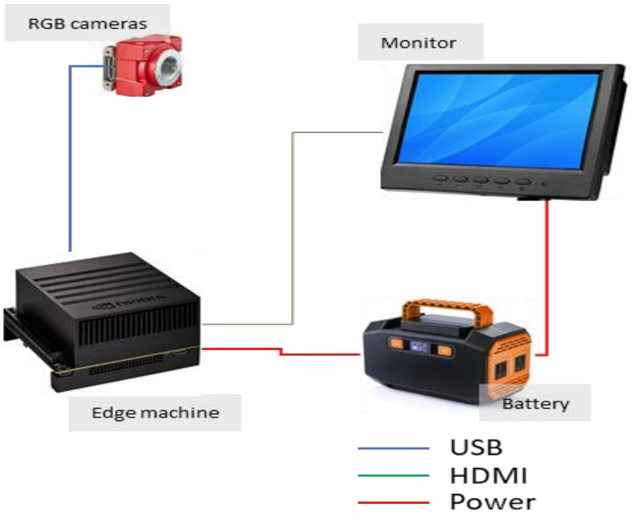


Figure 4: Experiment Design Using Edge Device. This Setup has been Placed on The Trolley

This meticulous approach serves a critical purpose by minimizing calibration errors that might otherwise result from environmental factors such as light deflection, refraction, and variations in camera placement precision. Traditional geometric methods traditionally used for determining 3D object positions are intentionally avoided due to their susceptibility to sinusoidal and tangential functions. The camera calibration procedure is executed by capturing images of the object at varying distances from the camera. This systematic approach enables us to establish a robust correlation between object distances and the corresponding shifts in the object's center.

Visual SLAM

Real-Time Appearance-Based Mapping (RTAB-Map, Labbe et al. (2014)) represents a visual graph SLAM algorithm founded on closure detection principles. The detector leverages a bag-of-words algorithm to assess the similarity between new camera images and those captured at previously visited locations. Each such closure event introduces a new edge into the camera position graph, followed by graph optimization. RTAB-Map8 offers a user-friendly graphical interface that presents the results of the loop closure detector (Fig. 12a), visual odometry, and point cloud (Fig. 12b). Furthermore, it provides the capability to filter and export point clouds in PLY or PCD formats. RTAB-Map seamlessly accommodates various camera types, including Kinect RGB-D cameras, ZED stereo cameras, conventional stereo pairs, and more.

Real-Time Appearance-Based Mapping (RTAB-Map, Labbe et al. (2014)) is a SLAM (Simultaneous Localization and Mapping) algorithm. SLAM is a fundamental problem in robotics and computer vision. It involves a system, like a robot or a camera, moving through an unknown environment while simultaneously creating a map of that environment and determining its own position within it. One of RTAB-Map's standout features is loop closure detection. When the prototype system revisits a location it has seen before, RTAB-Map recognizes this event. This recognition is crucial because it helps correct errors that may accumulate over time, enhancing the overall accuracy of the generated maps. Loop closure involves detecting and closing loops in the path of the system.

RTAB-Map can fuse data from various sensors, such as cameras (both RGB and depth), laser scanners, and inertial measurement units (IMUs), in our case, we used three cameras as input

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008
to RTAB-Map. This multisensory data fusion allows for a more comprehensive understanding of the environment and helps compensate for limitations in individual sensor types. It employs a graph-based optimization technique. In the SLAM problem, the system maintains a graph that represents its path and the relative positions of landmarks, which the features in the environment. RTAB-Map optimizes this graph to refine the estimated positions and improve map accuracy.

Integration with ROS (Robot Operating System) is facilitated through the RTAB-Map *rospkg* package, providing seamless compatibility. The outcomes of the algorithm encompass a 3D map presented as a point cloud, visual odometry data, and a 2D map represented as an occupancy grid map. In this grid map, each cell signifies the likelihood of space occupancy. To visualize and interact with this data, the standard *RViz* package is employed.

2D to 3D Object Localization and Mapping

Once the mapping is complete, the system now has the 3D coordinates of the detected object within the environment. 2D coordinates (bounding box) of the detected object and, with the help of additional sensor data (such as depth information from a depth camera), calculates the object's position in a 3D map. This information mapped to the 3D space, where the z-axis represented the height of the tree, *x* and *y* location from bounding box and ID of each of the tree.

Results and Discussion

For every detected object's position, there is a corresponding collection of images. These images are captured with the object positioned in slightly different ways in relation to the camera's lens. The purpose of this adjustment is to guarantee that the results obtained from these images are precise and dependable. By capturing the object from various angles or positions, any potential errors or uncertainties in the object's location can be minimized or corrected, resulting in more accurate and reliable data.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008

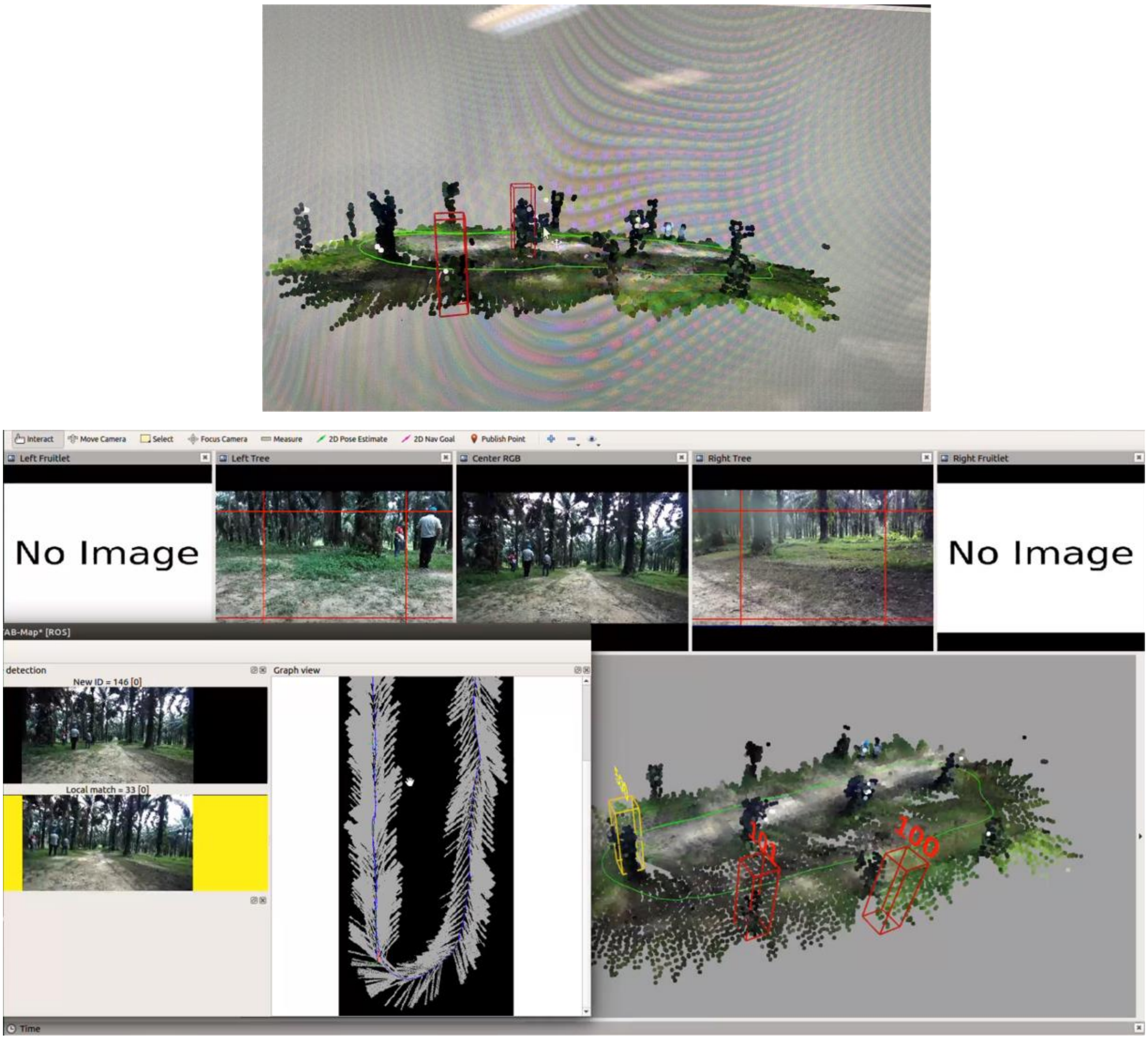


Figure 5: Results of The Detected Object Mapped into The 3D Space Using RTAB-Map

As depicted in Figure 5, it becomes evident that the total count of detected trees perfectly matches the number of trees that have been tagged. The color-coded bounding boxes play a pivotal role in visual representation. A red bounding box signifies the left tree's detection by the left camera, while a yellow bounding box indicates the tree's detection by the right camera.

While our algorithm is engineered to cluster multiple detections of the same tree, it exhibits challenges in instances where the algorithm registers identical trees as distinct entities, even when their proximity meets the defined minimum distance for clustering.

Furthermore, an important consideration relates to the placement of text indicating the tree's identification. Ideally, this text should be positioned above the corresponding bounding box. However, during live detection, it becomes apparent that this feature does not perform as expected, necessitating further attention for resolution. Lastly, the specific video under examination has proven successful in achieving local matches and conducting loop closure detection. It's noteworthy that a significant portion of the matched features primarily originates from the ground area. These observations provide valuable insights into both the system's

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008
achievements and the challenges it faces in the processes of detection and localization. These insights serve as a foundation for ongoing refinement and optimization efforts."

Within the RTAB-Map interface, a graph view represents the specific pathway within the plantation area. As RTAB-Map registers new images, they are marked with a yellow background (Figure 5). When matches confirming the location are detected, the current position (marked as red circle in Figure 6) transitions to green as illustrated in Figure 7, while Figure 6 shows the RTAB-Map still searching for local point, thus the path is deviated from the initial position (marked as yellow circle in Figure 6). Subsequently, RTAB-Map adjusts and refines the path, as illustrated in graph view in Figure 7.

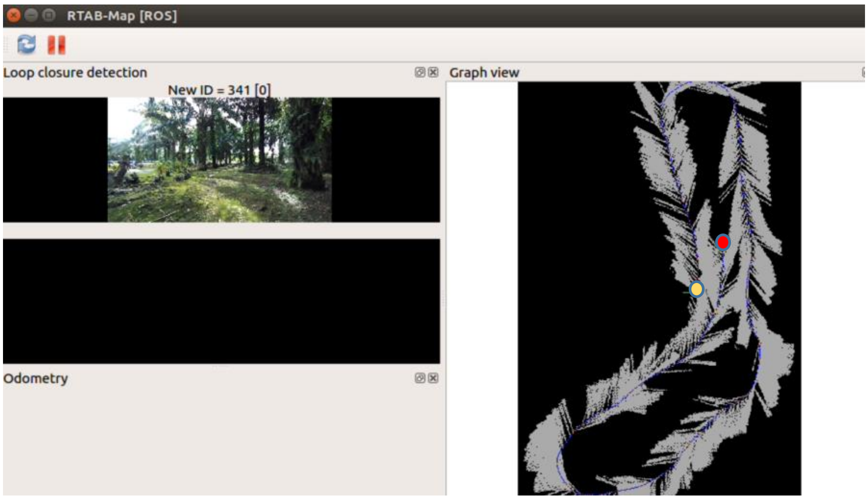


Figure 6: RTAB-Map Still Searching for Local Point

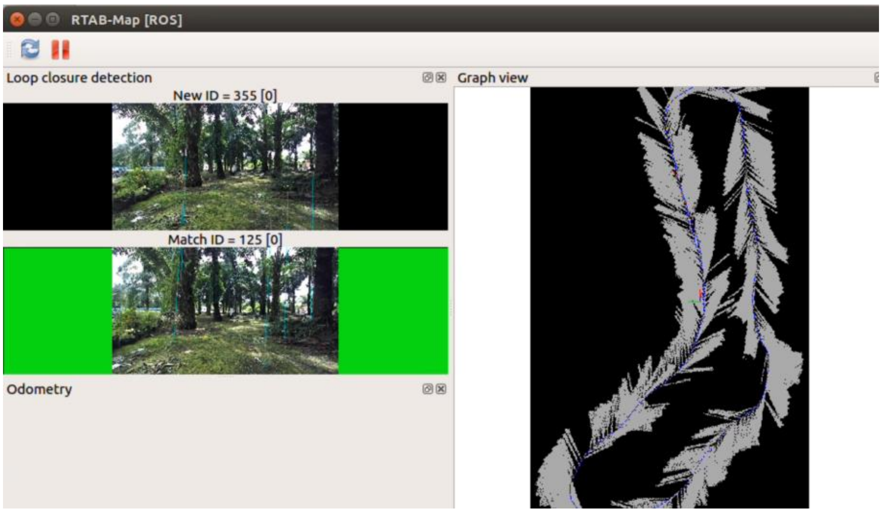


Figure 7: RTAB-Map Found The Match ID for Loop Closure Detection

Conclusion

This paper presents a novel approach for real-time 3D mapping and localization using tree and fruitlet detection data. The study leverages visual SLAM with RTAB-Map to create an accurate 3D map of tree locations within a real palm oil plantation. The experiments were conducted using a custom-designed UAV prototype equipped with ZED cameras, an edge device, and a monitor. The preliminary results demonstrate the promising accuracy of the 3D map; however,

Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008

there remains room for quantitative accuracy enhancement in future experiments. Moreover, the proposed system will enable the real-time transmission of harvest-ready location data to plantation owners via an IoT system for future works. This research offers valuable assistance to plantation management, resulting in significant time and labor cost savings.

Acknowledgment
This work was carried out with the financial support from the Ministry of Higher Education of Malaysia under the research grant GGPM-2023-023.

References

Castiblanco, J. S., & Mosquera, M. (2010). Evaluación económica de dos sistemas de evacuación de fruta: alce manual y cable vía. Rev. Palmas, 31(1), 33-42.

Daud, M. M., Kadim, Z., & Woon, H. H., (2023) "Detection of Oil Palm Tree and Loose Fruitlets for Fresh Fruit Bunch's Ready-to-Harvest Prediction via Deep Learning Approach," in IAENG International Journal of Computer Science.

Daud, M. M., Kadim, Z., & Woon, H. H. (2022, November). Loose Fruitlet and Fresh Fruit Bunch Detection for Palm Oil Harvest Management. In 2022 IEEE International Conference on Internet of Things and Intelligence Systems (IoTaIS) (pp. 13-18). IEEE.

Department of Occupational Safety and Health, Malaysia. "Occupational Accident and Statistics". [Accessed on 27 April 2021] https://www.dosh.gov.my/index.php/statistic-v/occupational-accident-statistics

Diao, Y. B. and B. Z. and N. X. and J. Z. and J. S. and Z. (2023). Vision-based navigation and guidance for agricultural autonomous vehicles and robots: A review. *Computers and Electronics in Agriculture, 205*(January). https://doi.org/10.1016/j.compag.2022.107584

Hannah, R., & Max, R. (2022). Forests and Deforestation. Retrieved January 3, 2022, from https://ourworldindata.org/forests-and-deforestation.

Hu, Y., Lu, Y., Xu, R., Xie, W., Chen, S., & Wang, Y. (2023). Collaboration Helps Camera Overtake LiDAR in 3D Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 9243-9252.

Junos, M. H., Khairuddin, A. S. M., Thannirmalai, S., & Dahari, M. (2021). An optimized YOLO-based object detection model for crop harvesting system. IET Image Processing, 15(9), 2112-2125.

Labbe, M., & Michaud, F. (2014). Online Global Loop Closure Detection for Large-Scale Multi-Session Graph-Based SLAM. In Proc. IEEE/RSJ Intelligent Robots and Systems (IROS), 2661–2666.

Lee, J. S. H., Ghazoul, J., Obidzinski, K., & Koh, L. P. (2014). Oil palm smallholder yields and incomes constrained by harvesting practices and type of smallholder management in Indonesia. Agronomy for Sustainable Development, 34(2), 501–513. https://doi.org/10.1007/s13593-013-0159-4.

Malaysian Palm Oil Council (2022). Monthly Palm Oil Trade Statistics 2020. Retrieved January 4, 2022, from https://mpoc.org.my/monthly-palm-oil-trade-statistics-2020/.

Mosquera, M., Ruiz, E., Fontanilla, C. A., Valderrama, M., López, D. F., Castro, L. E., & González, M. A. (2017). Costos de producción para el cultivo de la palma de aceite en Colombia. In Y. Moreno & E. Mantilla (Eds.), Mejores prácticas agroindustriales del cultivo de la palma de aceite en Colombia (pp. 427-435). Cenipalma.

Parveez, G. K. A., et al. (2021). Oil palm economic performance in Malaysia and R&D progress in 2020. J. Oil Palm Res, 33(2).

Journal of Information System and Technology Management
EISSN: 0128-1666
JISTM
Volume 8 Issue 33 (December 2023) PP. 91-101
DOI: 10.35631/JISTM.833008
Ruiz, E., Mesa, E., Mosquera, M., & Barrientos, J. (2017). Technological factors associated with oil palm yield gaps in the Central region in Colombia. Agron. Colomb., 35(2), 256-264.

Sofiah, S., Mohd, B., Sarah, S., Sallah, B., Azmeer, N., Ming, L., Aiman, Q., Tajul, B., Bin, B., Jailani, M., & Woon, H. H. (2023). Visual-based and Lidar-based SLAM Study for Outdoor Environment. 13–16.

Zheng, J., Fu, H., Li, W., Wu, W., Yu, L., Yuan, S., Tao, W. Y. W., Pang, T. K., & Kanniah, K. D. (2021). Growing status observation for oil palm trees using Unmanned Aerial Vehicle (UAV) images. ISPRS Journal of Photogrammetry and Remote Sensing, 173(January), 95–121. https://doi.org/10.1016/j.isprsjprs.2021.01.008.