International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0183

e ISSN: 2584-2854
Volume: 03
Issue:04 April 2025
Page No: 1122-1128

Tomato Harvesting Robot Using Yolov8, Midas and Inverse Kinematics for Smart Farming Automation

V. Muthupandi[1], *K. Kannan*[2], *J. Saroj Kanna*[3], *K. Esakki Bala Karthik*[4]

[1,3,4]*UG Scholar, Dept. of Mechatronics Engineering, Kamaraj College of Engineering and Technology, Virudhunagar, Tamil Nadu, India.*

[2]*Faculty Dept. of Mechatronics Engineering, Kamaraj College of Engineering and Technology, Virudhunagar, Tamil Nadu, India.*

Email ID: *vmuthupandi892@gmail.com*[1], *kannan_kcet@yahoo.co.in*[2] *jsarojkanna@gmail.com*[3], *balakarthik0103@gmail.com*[4]

Abstract

This project presents a cutting-edge tomato harvesting robot that combines advanced image processing and robotic manipulation to optimize the harvesting process. Manual tomato harvesting is labour-intensive and requires a significant workforce, especially during peak seasons. Labor shortages and rising wages increase production costs, making automation a more viable alternative. Human pickers may vary in their selection criteria, leading to inconsistent quality and harvesting of unripe tomatoes. Improper handling can cause bruising or mechanical damage, reducing the market value of tomatoes. Farmers struggle to track and optimize harvesting schedules, often leading to wastage or missed ripe tomatoes. An automated system with deep learning-based classification ensures uniform selection based on ripeness. A robotic system with a customized soft end-effector minimizes damage by ensuring gentle gripping and precise detachment. This robot uses a YOLOv8 model to detect tomatoes and determine if they are ripe, unripe or partially ripe. MiDaS depth estimation helps measure the exact position of tomatoes in 3D space for precise picking. A 6-DOF robotic arm moves towards the detected tomato using inverse kinematics for accuracy. A soft-gripper gently holds the tomato, while a scissor detaches it from the vine without damage. The robot moves on a track-based system, ensuring stable and smooth navigation in farms and greenhouses. The results demonstrate that the autonomous tomato harvesting robot is a feasible and effective solution for modern agriculture. By combining deep learning, computer vision and robotic automation, this system enhances yield efficiency, reduces post-harvest losses and promotes sustainable farming.

Keywords: *Tomato harvesting, YOLOv8, MiDaS, Inverse Kinematics, Track system, Robotic grippers.*

1. Introduction

Agriculture is a vital industry facing challenges such as labor shortages, rising costs, harvesting inefficiencies and post-harvest crop losses. Among various crops, tomatoes require careful handling due to their delicate nature, making manual harvesting labor-intensive and prone to inefficiencies. To address these challenges, this project focuses on developing an intelligent tomato harvesting robot that leverages advanced robotics and artificial intelligence to enhance productivity and precision. The robot employs high-resolution cameras and sophisticated image processing algorithms to detect and classify ripe tomatoes. The system analyzes multiple parameters such as color, texture and size to ensure only mature, healthy tomatoes are harvested. Deep learning models trained on large datasets enable the robot to improve its accuracy over time, ensuring optimal picking decisions. Instead of traditional wheeled or legged locomotion, the robot utilizes a track-based mobility system, which provides stable movement across uneven or muddy terrain, better traction in dense vegetation and improved adaptability to different field conditions. This enhances reliability, allowing the robot to operate effectively in both greenhouse and open-field environments without requiring complex terrain

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0183
e ISSN: 2584-2854
Volume: 03
Issue:04 April 2025
Page No: 1122-1128

mapping. A robotic arm equipped with soft, adaptive grippers ensures damage-free tomato harvesting. The arm is designed with multiple degrees of freedom for precise movement and automated positioning to efficiently pick tomatoes from various plant heights and angles. Its human-like dexterity enables selective harvesting while minimizing contact with non-target areas. The robot features real-time decision-making capabilities, allowing it to adapt to changing environmental conditions, prioritize harvesting based on ripeness levels and optimize picking sequences. It can also communicate with a central system for remote monitoring and control. These intelligent features maximize efficiency, reduce waste and enhance overall productivity. To further optimize the harvesting process, the robot can be integrated into smart farming ecosystems through IoT connectivity for real-time data transmission, cloud-based analytics to improve decision-making and AI-driven predictive maintenance for uninterrupted operation. These advancements make the system scalable and adaptable for different agricultural setups. By automating the harvesting process, the tomato harvesting robot reduces labor dependency, increases efficiency and minimizes crop losses. It ensures only ripe, healthy tomatoes are harvested, improving quality control and market value. The system also supports sustainable farming practices by lowering operational costs and optimizing resource utilization. Through the integration of robotics, AI and advanced image processing, this project presents a transformative solution for modern agriculture, paving the way for more precise, efficient and sustainable harvesting practices. [1]

2. Related Work

There are two types of research related to our project carried out by many researchers. Those are robotic manipulation and tomato detection.

2.1. Robotic Manipulation

Shiu et al. propose a system for autonomous fruit harvesting using a robotic arm, integrating a vision-based approach with a manipulation model. The system uses RGB-D cameras to detect ripe fruits and estimate their positions accurately. The manipulation model guides the robot arm's trajectory for fruit harvesting. Experimental results show the system can harvest fruits effectively without causing damage [1]. Kim et al. introduce a robotic arm equipped with a vision system for autonomous tomato harvesting, combining a vision system with a picking algorithm to identify ripe tomatoes and remove them from vines. The system's accuracy and ability to handle the fruit carefully were demonstrated in experiments, showing great promise for future automation in agriculture [2]. Zhao et al. explore a multi-modal approach to autonomous tomato harvesting, combining RGB images and 3D depth data to detect ripe tomatoes in cluttered environments. The system, integrated with a robotic arm, can pick tomatoes with high precision and efficiency, reducing manual labor and improving the harvesting process. Jun et al. describe a tomato harvesting robot that integrates 3D perception, manipulation and an end-effector to automate the harvesting process. The system's effectiveness is demonstrated through testing, showing its capability for autonomous and efficient tomato harvesting [9].

2.2. Tomato Detection

Wu et al. developed a vision-guided robotic system for tomato harvesting that uses machine learning algorithms to identify ripe tomatoes and robotic arms to pick them. The system uses support vector machines (SVM) and convolutional neural networks (CNN) to process images of tomatoes, achieving high classification accuracy. This system performs excellently in controlled environments and real-world agricultural settings, offering a promising solution for automating the tomato harvesting process while minimizing damage [3]. Liu et al. proposed a deep learning-based approach for classifying the ripeness of tomatoes using image processing techniques. They used convolutional neural networks (CNN) to analyze the color and texture of tomatoes, determining ripeness levels with high accuracy. This model works well in varying environmental conditions, reducing waste and improving efficiency in agricultural operations [4]. Yang et al. presented a deep learning-based system for real-time detection of ripe tomatoes, training models on a large dataset of tomato images. This system can accurately detect ripe tomatoes under varying lighting and background conditions,

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0183
e ISSN: 2584-2854
Volume: 03
Issue:04 April 2025
Page No: 1122-1128

integrating with a robotic manipulator to autonomously harvest the tomatoes [5]. Lee et al. investigated the use of a hybrid deep-learning model for real-time tomato detection in agricultural settings, combining traditional image processing techniques with deep neural networks [6]. Song et al. proposed TDPPL-Net, a lightweight real-time tomato detection and localization system optimized for low-cost industrial personal computers (IPCs). The system achieves an impressive 93.36 Mean Average Precision (mAP) accuracy and 31.41 FPS speed, making it suitable for complex agricultural environments [7]. Huang et al. introduced a fuzzy Mask R-CNN model for automatically identifying the ripeness of cherry tomatoes, achieving 98% accuracy in predicting ripeness [8].

2.3. Hardware Components

The tomato harvesting robot is powered by a Raspberry Pi with 8GB RAM, ensuring smooth processing for real-time operations. It uses an RGB camera module for tomato detection and depth estimation. The L298N motor driver controls four 12V DC motors for mobility, while six MG995 metal servos operate the robotic arm for precise tomato picking. The system integrates computer vision with YOLOv8 to detect ripe tomatoes and execute accurate grasping, making the harvesting process efficient and autonomous. [2]

2.4. Software Components

The tomato harvesting robot operates using Raspberry Pi OS, with YOLOv8 for real-time tomato detection and MiDaS for depth estimation. OpenCV is used for image processing, while TensorFlow supports deep learning tasks. Scikit-learn helps with any additional machine learning-based classification. pigpio and PCA9685 manage servo control, while PWM & GPIO libraries handle motor operations. The system integrates deep learning, image processing and motor control to enable precise and efficient tomato harvesting. [3]

2.5. Simulation

The simulation of a tomato harvesting robot in ROS has been successfully implemented, integrating multiple components for a complete system. Using URDF and Xacro, the robot's mechanical structure, including its robotic arm and end-effector, has been modeled. In Gazebo, a virtual farm environment with tomato plants has been created for realistic testing. For perception, YOLOv8 has been integrated for real-time tomato detection, along with MiDaS for depth estimation, enabling precise grasping. Motion planning has been implemented using MoveIt! to efficiently control the robotic arm's movements. Additionally, ROS nodes handle control and navigation, ensuring smooth operation. This simulation provides a reliable platform for testing and refining the robot's performance before real-world deployment. The 3D model of the robot shown in (Figure 1)

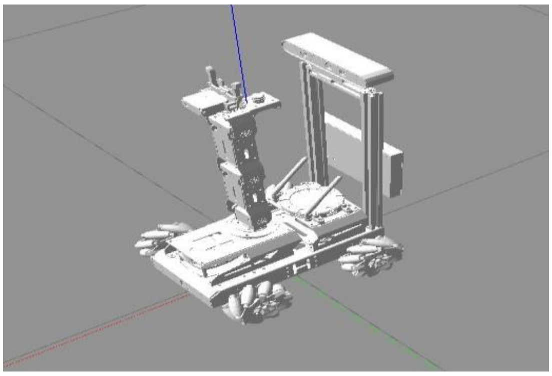

Figure 1 **Robot 3D Model**

2.6. Environment Setup

The glasshouse serves as the main structure where the robot operates, providing a controlled environment for tomato cultivation. Inside the glasshouse, individual tomato plants are placed in tomato plant pots, simulating real farming conditions. In the Gazebo framework, the "world" refers to a comprehensive environment encompassing objects, global parameters and physics features, with default settings predefined by Gazebo. This world consists of both static and dynamic elements. Static objects, such as support frames, lighting systems and greenhouse structures, are defined by their visual and collision geometry. In contrast, dynamic objects, including robotic harvesters and conveyor systems, not only have visual and collision geometry but also possess inertia information, which influences their mobility and interactions within the environment. FIG 2 represents the glasshouse for the simulation environment. (Figure 2) [4]

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0183
e ISSN: 2584-2854
Volume: 03
Issue:04 April 2025
Page No: 1122-1128

Figure 2 **Glasshouse in Gazebo**

2.7. Tomato Model

Each tomato plant consists of several components: the stem, which serves as the main structure supporting the leaves; the leaves, which are modeled to resemble real tomato plant foliage for accurate simulation; and the fruits, representing tomatoes at different growth stages, as shown in Figure 1. In the Gazebo universe, objects can be created using various methods, including 3D modeling tools, conventional geometric shapes, or importing from the model database. The environment consists of various objects, including the greenhouse landscape, supporting stands, tomato plants at different growth stages, ripe and unripe tomatoes, stems and leaves, all contributing to a realistic farming simulation. FIG 3 represents the tomato leaves and FIG 4 represents the tomato design. (Figure 3)

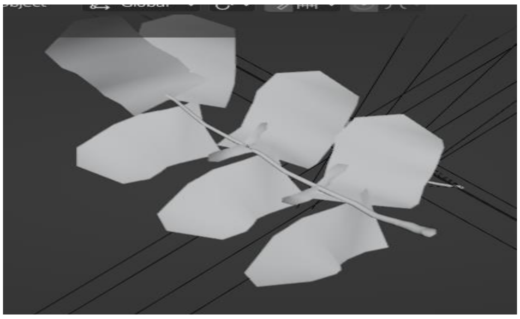

Figure 3 **Tomato Leaves**

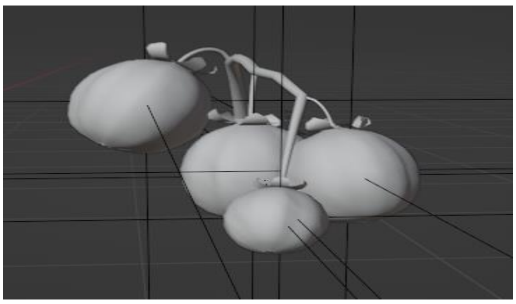

Figure 4 **Tomato Design in Blender**

3. Methodology

The robot's working has split into three major sections. These are Track Based Movement System, Tomato Detection, Grasping of Tomato from the plant.

3.1. Track-Based Movement System

The tomato harvesting robot operates on a track-based locomotion system to ensure precise navigation within the controlled agricultural environment. Instead of conventional autonomous navigation, the system follows a predefined track, guaranteeing systematic traversal along crop rows. The track-based mechanism eliminates the need for complex localization algorithms, thereby reducing computational overhead and enhancing operational reliability. Movement is actuated using high-torque DC motors equipped with quadrature encoders for real-time velocity and position feedback. An electromechanical control unit regulates motor speed and ensures synchronized movement. The system optimizes power efficiency, minimizes mechanical drift and facilitates consistent harvesting cycles. Block diagram of the proposed system shown in (Figure 5) [5]

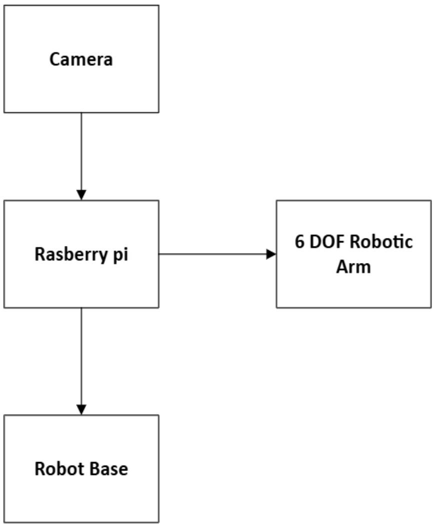

Figure 5 **Block Diagram of the Proposed System**

3.2. Tomato Detection Through Deep Learning

The tomato detection module leverages YOLOv8, a state-of-the-art convolutional neural network

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0183
e ISSN: 2584-2854
Volume: 03
Issue:04 April 2025
Page No: 1122-1128

optimized for real-time agricultural vision tasks. The robot's high-resolution RGB camera module captures images, which are processed through YOLOv8 for precise tomato localization, ripeness estimation and anomaly detection. Instead of traditional depth-sensing LiDAR systems, the robot employs MiDaS depth prediction, which reconstructs 3D spatial information from monocular RGB inputs, enabling accurate fruit positioning. The data pipeline consists of image acquisition, annotation preprocessing, dataset augmentation and neural network training. Images are labelled in YOLO format with bounding box coordinates defining tomato positions. The trained model operates efficiently on the Raspberry Pi 8GB RAM, utilizing TensorFlow Lite and ONNX optimizations for low-latency inference. The system enhances automated harvesting efficiency by identifying fruit maturity levels, disease symptoms and occluded objects, improving yield estimation and quality control. Tomato detection is shown in (Figure 6)

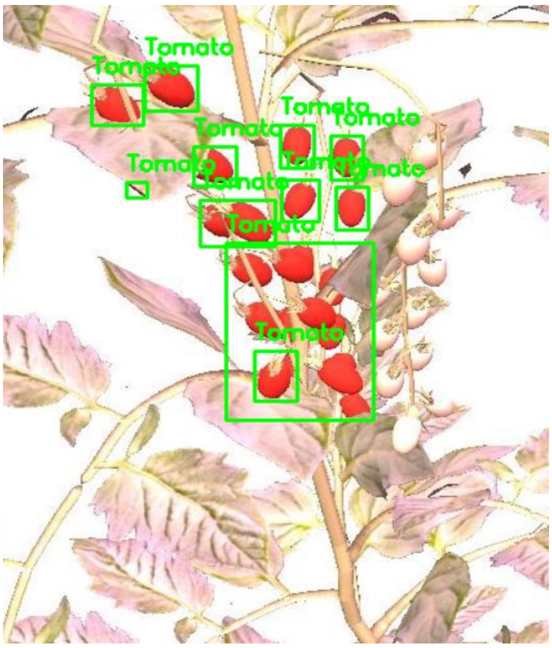

Figure 6 **Tomato Detection**

3.3. Tomato Grasping with Soft Gripper

Tomato grasping is executed by a 6-DOF robotic manipulator controlled via six MG995 metal-geared servo motors. The grasping framework integrates deep learning-based object detection for tomato localization, while MiDaS depth estimation provides real-time 3D spatial coordinates. Inverse kinematics algorithms compute the optimal joint configurations to align the soft robotic gripper with the detected fruit. The grasping workflow follows sequential stages: approach trajectory planning, pre-grasp adjustment, adaptive grasp execution and force-feedback validation. A pneumatic or elastomer-based soft gripper ensures delicate handling, mitigating potential fruit bruising. Real-time proprioceptive feedback from force sensors and vision-based slip detection allows dynamic grip force modulation, ensuring precise harvesting. (Figure 7) [7]

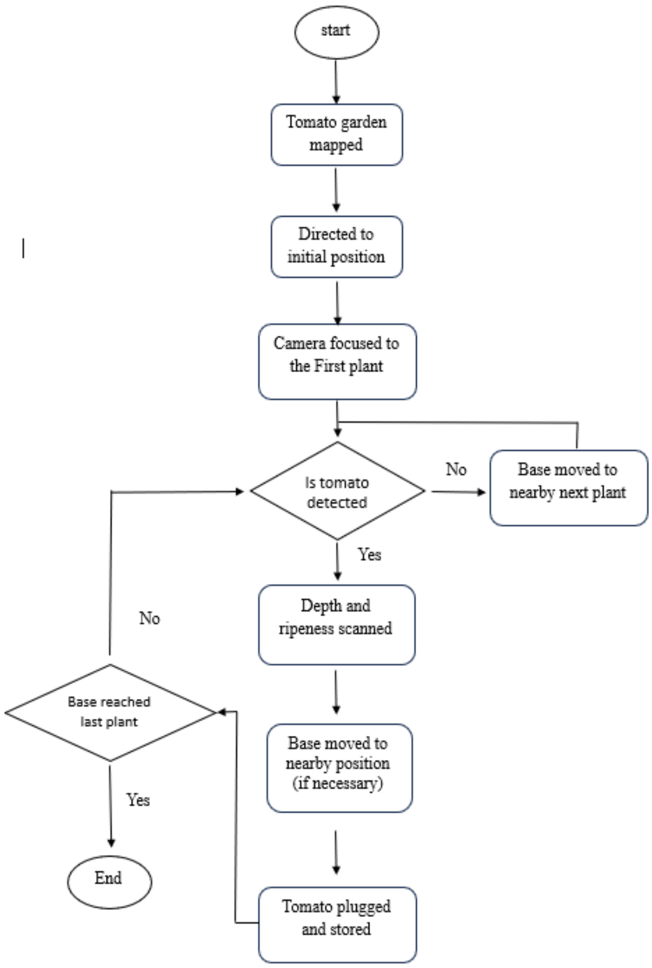

Figure 7 **Work Flow**

The fusion of computer vision, real-time control systems and intelligent grasp planning enhances adaptability in unstructured farming environments, leading to superior harvesting accuracy and minimal produce damage. Work flow of the proposed system shown in (Figure 7) [8]

4. Results and Discussion

4.1. Results

The tomato detection system, developed using deep learning techniques, was tested in controlled greenhouse environments to accurately detect both ripe and unripe tomatoes. The system achieved a precision of 92%, a recall of 89% and a mean average precision of 90.5%. The detection was tested under different lighting and environmental conditions, with minimal effect on detection performance due to the robust training dataset.

4.2. Discussion

The system also performed well in the real-time processing of Video streams, achieving an inference rate of approximately 25 frames per second (FPS) on a Raspberry Pi. The tomato detection project demonstrates significant advancements in autonomous agricultural robots using deep learning. However, there are several limitations, such as occlusion issues, overlapping fruits and training dataset limitations. Future work could include integrating additional sensors, enhancing data augmentation, fine-tuning the detection model with a more diverse dataset and implementing a hybrid model combining object detection and semantic segmentation. By continuing to enhance the detection system and integrating advanced multi-modal sensor fusion, autonomous robotic harvesting could be seamlessly realized for large-scale commercial farming.

Conclusion

This project successfully addresses the challenges associated with tomato harvesting, including labour shortages, inefficiencies and the need for gentle handling of delicate fruits. The track-based mobility system provides stable and efficient movement across various field conditions, eliminating the limitations of traditional wheeled platforms. The system adapts to changing environmental conditions by integrating real-time decision-making algorithms, further optimizing performance. These results confirm that the proposed solution is a viable and scalable approach to addressing the limitations of manual tomato harvesting. Future work may focus on improving system robustness, expanding adaptability to different crop types and integrating AI-driven predictive maintenance for enhanced reliability. With further advancements, this technology can play a crucial role in modernizing agricultural practices, making them more efficient, sustainable and economically viable. [9]

Acknowledgements

We express our sincere gratitude to all individuals and organizations that contributed to the successful completion of this project. We extend our deepest appreciation to our mentors and professors for their invaluable guidance, constructive feedback and continuous support throughout the development of this work. We would also like to acknowledge the support received from our college for providing the necessary resources, infrastructure and technical assistance that enabled us to carry out this research effectively. Special thanks to our peers and colleagues for their insightful discussions and collaborative efforts, which played a crucial role in refining our approach.

References

[1]. R. Singh, L. Seneviratne and I. Hussain, “A deep learning-based approach to strawberry grasping using a telescopic-link differential drive mobile robot in ROS-Gazebo for greenhouse digital twin environments,” IEEE Access, vol. 13, pp. 361–381, 2025, doi: 10.1109/ACCESS.2024.3520233.

[2]. N. Xu, Z. Li, J. Kang, Q. Meng and M. Niu, “Agricultural vehicle automatic navigation positioning and obstacle avoidance technology based on ICP,” IEEE Access, vol. 12, pp. 85940-85954, May 2024, doi: 10.1109/ACCESS.2024.3402743.

[3]. Z. Wang, P. Li, Z. Wang and Z. Li, "APG-RRT: Sampling-Based Path Planning Method for Small Autonomous Vehicle in Closed Scenarios," IEEE Access, vol. 12, pp. 25731-25739, Feb. 2024, doi: 10.1109/ACCESS.2024.3359643.

[4]. J. Pak, J. Kim, Y. Park and H. I. Son, "Field Evaluation of Path-Planning Algorithms for Autonomous Mobile Robot in Smart Farms," IEEE Access, vol. 10, pp. 60253-60263, Jun.

2022, doi: 10.1109/ACCESS.2022.3181131.

[5]. Y. Chen, B. Leighton, H. Zhu, X. Ke, S. Liu and L. Zhao, "Submap-Based Indoor Navigation System for the Fetch Robot," IEEE Access, vol. 8, pp. 81479-81488, Apr. 2020, doi: 10.1109/ACCESS.2020.2991465.

[6]. C. Song, K. Wang, C. Wang, Y. Tian, X. Wei, C. Li, Q. An and J. Song, "TDPPL-Net: A Lightweight Real-Time Tomato Detection and Picking Point Localization Model for Harvesting Robots," IEEE Access, vol. 11, pp. 37650-37662, Mar. 2023, doi: 10.1109/ACCESS.2023.3260222.

[7]. J. Jun, J. Kim, J. Seol, J. Kim and H. I. Son, "Towards an Efficient Tomato Harvesting Robot: 3D Perception, Manipulation and End-Effector," IEEE Access, vol. 9, pp. 17631-17640, Jan. 2021, doi: 10.1109/ACCESS.2021.3052240.

[8]. Y.-P. Huang, T.-H. Wang and H. Basanta, "Using Fuzzy Mask R-CNN Model to Automatically Identify Tomato Ripeness," IEEE Access, vol. 8, pp. 207672-207684, Nov. 2020, doi: 10.1109/ACCESS.2020.3038184.

[9]. N. Prieto-Fernández, S. Fernández-Blanco, Á. Fernández-Blanco, J. A. Benítez-Andrades, F. Carro-De-Lorenzo and C. Benavides, "Weighted Conformal LiDAR-Mapping for Structured SLAM," IEEE Trans. Instrument. Meas., vol. 72, pp. 8504110, 2023, doi: 10.1109/TIM.2023.3284143.