

Logistics Robots Based on SLAM: Current Status and Future Trend

Wenhao Li¹, Zuodong Liu^{2,4,*}, Zixuan Zhang³

¹Detroit Green Technology Institute, Hubei University of Technology, Hubei, China

²School of mechanical engineering, Tiangong University, Tianjin, China

³School of Mechatronic Engineering, Guangdong Polytechnic Normal University, Guangdong, China

⁴zuodongliu95@gmail.com

*corresponding author

Abstract. The application of logistics robots has become an important part of improving industrial production efficiency, and the key technology to achieve precise positioning and path planning of logistics robots is SLAM (simultaneous localization and mapping) technology. At present, this application has gradually become the focus of research by scholars National and international. Given the growing importance of this technology, this paper provides a comprehensive analysis and summarizes some of the recent literature relating to this technology. To gain a clearer understanding of the state of research and development in this area, this paper provides a comprehensive analysis and summary of some recent literature. Firstly, the paper carefully summarizes the application of SLAM technology in logistics robotics operations. Secondly the current status and development trends of logistics robots. Finally, the paper enumerates the pros and cons of the current technology advantages and limitations of the current technology and suggests possible solutions for future research in logistics robotics.

Keywords: SLAM, path planning, localization, visual slam.

1. Introduction

As science and technology continue to evolve, intelligent robots are gradually integrating into people's lives, and the great convenience they bring leads to increasing demand, and experts and scholars in various industries are actively promoting the development of the field of intelligent robots. There are many practical applications for robotics, including indoor cleaning robots, self-driving Vehicles in the field, underwater environmental detection robots, UAVs (Unmanned aerial vehicles) and even the increasingly popular virtual scenarios of AR and VR [1]. With the rapid evolution of the logistics industry and the rise of intelligent warehousing, movable logistics robots have been favored by many enterprises and scholars. However, in the absence of high accuracy localization and mapping technology, logistics Robots are not able to move autonomously indoors.

SLAM (Simultaneous localization and mapping) provides an effective method for this problem. It can provide spatial localization information based on its own position and build maps and virtual scenes. SLAM technology refers to simultaneous localization and map construction [2]. Localization means automatically determining the robot's position and surrounding objects in the world coordinate system,

while mapping means constructing a map of the robot's perceived environment. The robot is equipped with specific sensors for determining the robot's path, estimating the robot's movement, and constructing a map of the environment through observation of the environment without advance information about the surroundings.

The aim of this article is to provide theoretical support and practical guidance for the intelligent upgrading of the logistics industry by analyzing the current situation and development trend of logistics robot path planning technology based on the SLAM method. In order to understand SLAM technology and logistics robotic path planning technology research advancements, relevant national and international literature is reviewed, which provides theoretical support for this paper. This paper looking forward to the future development direction of SLAM technology in logistics robot path planning in light of the current technology development trend and market demand. This includes the trend of technological innovation, the expansion of application scenarios, and the possible challenges and opportunities. For the intelligent transformation of the logistics industry, the author expects useful insights and suggestions.

2. Slam on logistics robots

2.1. Environment Sensing and Localization Composition

The origin of Slam technology can be traced back as early as 1985, SLAM can be simply summarized as a problem of estimating the state of the mechanism itself, that is modeling the observed environmental data and eliminating the errors caused by sensors through appropriate optimization methods. Early researchers categorized SLAM techniques into vision SLAM and Lidar SLAM depending on the type of sensors used in the SLAM technique, such as cameras and Lidar [3]. For Laser Slam, the Lidar hardware continuously releases point cloud data at 10 HZ during the working process, the feature extraction module receives the data and extracts the feature points after preprocessing such as removing noise points. At this time, the localization module will look for the correspondence of the feature points, and output the localization and map building information with high frequency and low frequency respectively, and the combination of the two obtains the high-precision, real-time laser odometer. As for vision SLAM, the information in the environment, i.e., the surrounding images, is first acquired by binocular or RGB-D cameras in frame units. Subsequently, the feature point method is used to calculate the camera position by front-end Visual Odometry (VO) for feature point extraction and comparison.

2.2. Path Planning Algorithms

SLAM technology can help the mobile robot to obtain the surrounding map and information, but to move to the destination also needs the support of path planning algorithms.

The most frequently utilized global path planning algorithms for raster maps are Dijkstra's and A*. The Dijkstra algorithm represents a classic search algorithm, its underlying principle being that of breadth-first search. The algorithm employs breadth-first search to address the shortest path problem for a single source in an empowered directed or undirected graph. This process ultimately yields a shortest path tree. The algorithm is frequently employed in routing algorithms or as a component of other graph algorithms. The A* (A-Star) algorithm is a heuristic search algorithm that is effective for solving a variety of search problems. It is used in many applications, including indoor robot path search and game path search for animation. It is a type of graph search algorithm. It is a fusion algorithm based on depth-first search (DFS) and breadth-first search (BFS), which uses certain principles to determine which node to select. In recent years, alongside the expansion of the computer field and the ongoing pursuit of knowledge by researchers, there has been a proliferation of novel SLAM techniques. These include CNN-SLAM (based on semantic pattern recognition methods) and DeepVO (based on end-to-end deep learning methods), as well as composite SLAMs such as RTAB-MAP and VINS (IMU + vision). [4].

Currently, there are many mature algorithms for path planning in SLAM. Among them, A* is the more widely used algorithm. It is a heuristic search algorithm derived from the combination of Dijkstra's algorithm and the BFS algorithm. The principle of the A* algorithm is specifically that the goal is to

minimize the total cost, and when run, each expansion finds the most promising path node that reaches a predetermined location, so that it always plans a shortest path given the start and goal locations in a known map.

Ivan Maurovic et al. proposed an extended algorithm based on the D* algorithm that solves the shortest path problem with negative edge weights [5], effectively solving the problem of adding negative edge weight units to the algorithm that cannot be solved in the D*, A* algorithm to use the graph search algorithms for shortest paths that are common in robotics.

2.3. Application of SLAM on mobile robots

In recent years, with the growth of e-commerce, express delivery and other industries, people's requirements for logistics efficiency and automation are getting higher, and this promotes the application of SLAM in the field of logistics robotics.

At present SLAM is applied to mobile robots, mainly categorized into visual sensors and Lidar and multi-sensor fusion.

In 2007, Grisetti proposed the use of Gmapping particle filtering for robot's position estimation. Gmapping is usually used with lidar for 2D mapping of indoor environments, and is currently widely used for lidar mapping [6]. However, it is highly dependent on the odometer, and the particles cannot be repaired if they become noisy during propagation. In 2016, Google introduced Cartographer. Cartographer algorithm is an open source slam code algorithm based on a graphical optimization. Cartographer improves on Gmapping to support robots to move while updating the map, using graph optimization to solve for the position and reduce cumulative errors. Although the 2D lidar can work, it still has limitations, not being able to measure the height of the object, and difficult to complete the 3d map construction. So in 2014, Ji Zhang proposed the loam algorithm, which combines the data processing of lidar and visual odometry techniques to improve the accuracy of map building and positioning in 3D environments by optimizing point cloud matching [7].

Visual SLAM technology has also made tremendous progress and become more mature in the past few years. In 2007, Davison et al. proposed MonoSlam [8]. Monoslam is a monocular camera-based slam method, the estimation of camera motion and the construction of environment maps are realized by estimating and fitting feature points in continuous frames. Monoslam pioneered the research direction of using a Monoslam pioneered the research direction of using monocular camera for SLAM, even though it is easy to lose the feature points in the feature points, but it lays the foundation for the subsequent research and development of related technology. In the same year, Kelvin et al. proposed PTAM, using two main threads, tracking and mapping, and was The first system to employ nonlinear optimal techniques [9]. As Visual Slam technology is maturing, in 2014 Engel et al. proposed LSD-SLAM [10]. LSD-SLAM uses the direct method, directly utilizing pixel information in the image rather than relying on feature point extraction, providing an effective solution for localization building in large-scale environments. In 2015, R. Mur-Artal et al. proposed ORB-SLAM combining ORB feature point extraction and matching, closed-loop detection, and graph optimization, encompassing all the processes in the classical Visual SLAM framework (see Figure 1) [11].

ORB-SLAM greatly improves the stability and accuracy of localization mapping than the previous monocular SLAM techniques. One year later, R. Mur-Artal et al. proposed ORB-SLAM2 based on ORB-SLAM [12]. Compared to ORB-SLAM system is more flexible, supports monocular, stereo and RGB-D cameras, introduces a more efficient closed-loop detection mechanism, and optimizes the real-time performance aspect.

In the course of SLAM research, people also gradually realized that single sensors still have big limitations no matter how much they are enhanced, so multi-sensor fusion was proposed. In 2017, R. Mur-Artal et al. proposed Visual-Inertial SLAM in conjunction with ORB-SLAM, proposing a map reuse method and solving the most common problem of monocular cameras by combining visual and inertial sensors to solve the most common problem of scale ambiguity in monocular cameras [13]. Qin et al. proposed the VINS-Mono system, which also combines visual and inertial sensors, to obtain a highly accurate visual inertial odometer using a tightly-coupled, nonlinear optimization-based approach

[14]. The impact due to illumination variations and missing textures is reduced. In 2019, Jiang et al proposed a new SLAM framework based on graph optimization by combining Lidar sensors and vision sensors (see Figure 2) [15].

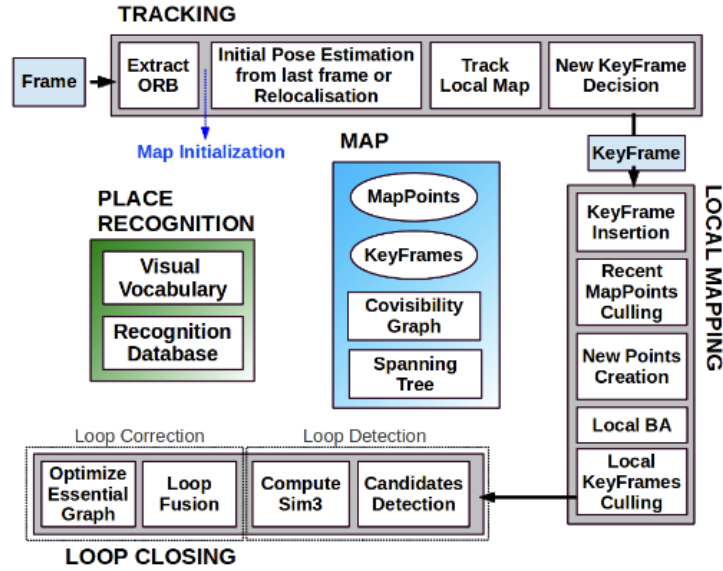


Figure 1. ORB-SLAM system overview

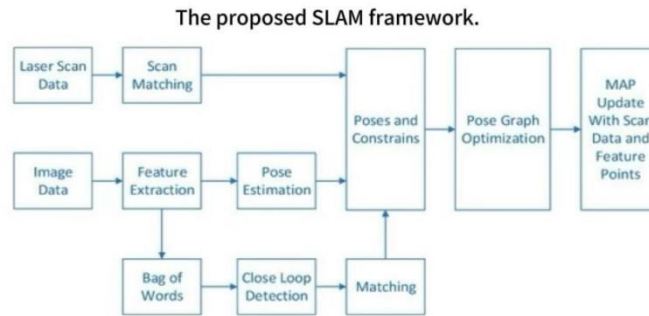


Figure 2. The proposed SLAM framework

In conclusion, the general direction and hotspot of current research is still Visual SLAM and Multi-Sensor Fusion.

2.4. Development trend of logistics robots based on SLAM methods

Visual SLAM usually uses a monocular camera, the application of logistics robots is less costly, and the monocular camera is suitable for most scenes and various tasks. However, the stability of vision sensors is general, especially within complex environments such as warehouses. visual sensors are able to provide environmental information in simple situations, but due to the high dependence of vision sensors on ambient lighting, they usually struggle to cope with low illumination or high dynamic range in the scene [16].

Since monocular cameras are difficult to achieve more accurate map construction, the use of multiple cameras to extract environmental information has also been proposed. For example, Danping Zou et al. mounted the cameras on different platforms so that the cameras can work together to construct global maps [17]. Although such a process allows for stable map construction in highly dynamic situations and more accurate map construction in static situations, the superposition of multiple cameras increases the data processing power and computational requirements of the system, resulting in higher costs.

Lidar SLAM has better stability than Visual SLAM, is not easily affected by ambient light, and has faster data processing speed more suitable for real-time applications. However, in a strongly reflective environment, it may cause the laser signal to be reflected back to the sensor, thus triggering erroneous distance measurements. Lidar has difficulty in acquiring semantic information about the environment as well as colors and boundaries relative to vision SLAM [18]. Unstructured Lidar points cannot render the scene texture, and low texture environments such as long corridors can be troublesome for Lidar SLAM [19]. And in the case of laser point cloud in a large number of dynamic objects, a large number of dynamic point cloud can lead to reimage trajectory results, which may override, overlap the static features in the scene, making the construction of maps inaccurate [20].

The classical single sensor for map construction has some limitations, especially within the complex environment of the warehouse, which is affected by light changes, dynamic obstacles, and environmental textures. Therefore, Multi-sensor fusion sensing technologies is the development trend of SLAM technology applied to logistics robots. Multi-sensor fusion can fully utilize the advantages of each sensor. The Lidar sensor gets the laser point cloud by scanning and then constructs the map by laser point cloud feature extraction, and the vision sensor constructs the map by camera image information, the vision sensor and the Lidar sensor complement each other's strengths, and work together to reduce the local uncertainty and improve the accuracy. In addition, the IMU contains accelerometers and gyroscopes, which help the robot to determine its position by measuring the acceleration and angular velocity. Multi-sensor fusion, enabling more diverse access to information about our environment and location, helping to enhance the localization accuracy. Therefore, multi-sensor fusion is the future development trend of logistics robots.

3. Discussion

3.1. Advantages of combining slam and logistics robots

Motion control and path planning of logistics robots have always been one of the difficult problems in research and industrial applications, and how to achieve the localization, environment feature extraction and map representation of logistics robots is the key to the effectiveness of robot. For example, multimodal intelligent logistics robots combine 3D CNN, LSTM and visual SLAM for route planning and control (see Figure 3).

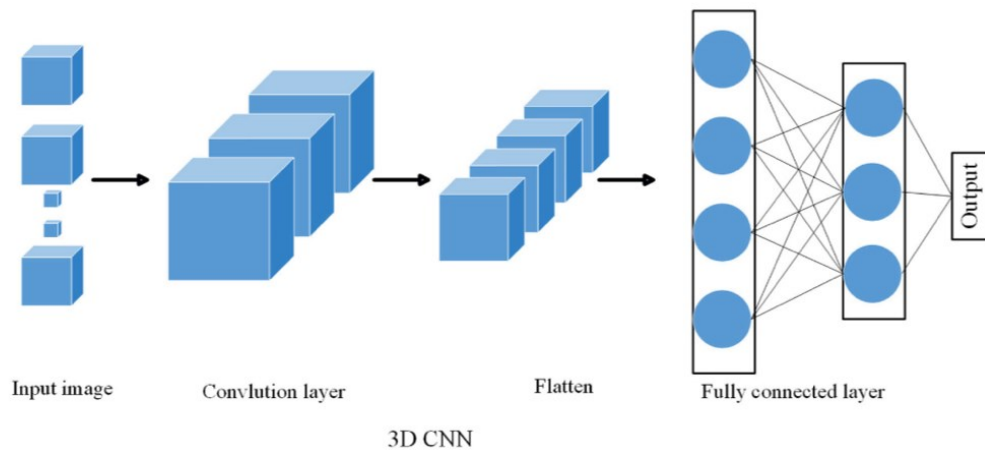


Figure 3. Flow chart of the 3D CNN mode.

This technique helps to improve the utility of logistics robots in complicated situations. The proposed approach combines object recognition, spatial temporal modelling and optimal route planning to enable logistics robots to navigate through complex situations with greater accuracy [21]. Firstly, the combination of SLAM and logistics robots to achieve path planning can greatly reduce human labour. Combined with slam technology, logistics robots can make autonomous decisions on real-time updated

maps and select the optimal path, which can improve transport efficiency while reducing human intervention and human labor. Secondly, the combination of SLAM and logistics robots for path planning reduces the need to rely on external infrastructure. Unlike robotic systems that rely on fixed base stations or external sensors, slam for logistics robots does not require additional infrastructure, but only relies on the robot's own sensors to operate. Finally, the combination of SLAM and logistics robots to achieve path planning can take advantage of the high precision positioning and navigation of SLAM technology. The use of SLAM technology allows the robot to achieve autonomous positioning in the warehouse. SLAM, or simultaneous localization and mapping, is a technology that allows a robot to estimate its own position while creating a map of its surroundings. This process provides basic data that can be used to plan the robot's journey. Robots can use advanced technologies such as LIDAR and sensors to accurately detect their surroundings and create detailed maps in real time. Next, an implementation of deep vision computing techniques is proposed for the recognition and analysis of objects in the warehouse. Optimization of the robot paths has been achieved using SORF and an adaptive weighting layer for aggregating the robot's position-sensitive regions has been created. Finally, the tasks of improving image quality and determining the optimal path have been successfully accomplished. An analysis of the experimental results checks the total length, obstacle identification and efficiency improvements achieved by this approach [22].

3.2. Limitations of combining slam with logistics robots

Although SLAM is currently widely used in logistics robots, there are still limitations. Firstly, SLAM technology's ability to build maps and process sensor data in real time is backed up by strong computing power, which puts demands on the processor and battery life of logistics robots. Secondly, slam has poor environmental adaptability and high dependence on sensors, and once the sensors are affected by ambient light, localization and map construction may be affected. And high quality sensors will increase the system cost. Thirdly, SLAM is subject to errors when faced with highly complex environments, such as those encountering highly reflective surfaces, crowded spaces, or large numbers of dynamic objects. Map construction and path planning may become inaccurate as a result. Moreover, in highly dynamic environments, slam may have difficulty processing data from all sensors in real time, resulting in delays.

3.3. Solutions

For a large amount of data processing and technical calculations, the SLAM algorithm can be optimized to reduce the complexity of the compute process and adopt more efficient algorithms. For the poor environmental adaptability of SLAM, it can be solved through multi-sensor fusion. Through the combination of laser sensors, visual sensors and IMU, the advantages complement each other. It can not only use laser point clouds and image information to build maps more accurately, but also use IMU to determine their own position to improve the accuracy of positioning. In addition, after reviewing and comparing the two most representative open source algorithms to solve SLAM, the ORB-SLAM2 and RTAB-Map methods were selected. These algorithms were used for handheld mapping experiments in which the sensors were manually moved around the environment to be mapped to facilitate the visual odometry. Additional experiments were conducted using the ORB-SLAM2 algorithm, which is based on the lateral alignment of the cameras in the Rob SLAM system [see *Figure. 4.*] and their remote manoeuvring on a predefined trajectory. Due to the lack of texture in the partitions and panels, optical odometry was lost in previously undetected areas of the environment. Any failure of algorithm execution in this case can be solved by employing a specific tessellation pattern in the spacer with different orientations. This study focuses on the research of multi-dimensional localization problems in uncharted environments or SLAM, especially the most advanced open source algorithms that can be used for three-dimensional environment modeling [23].

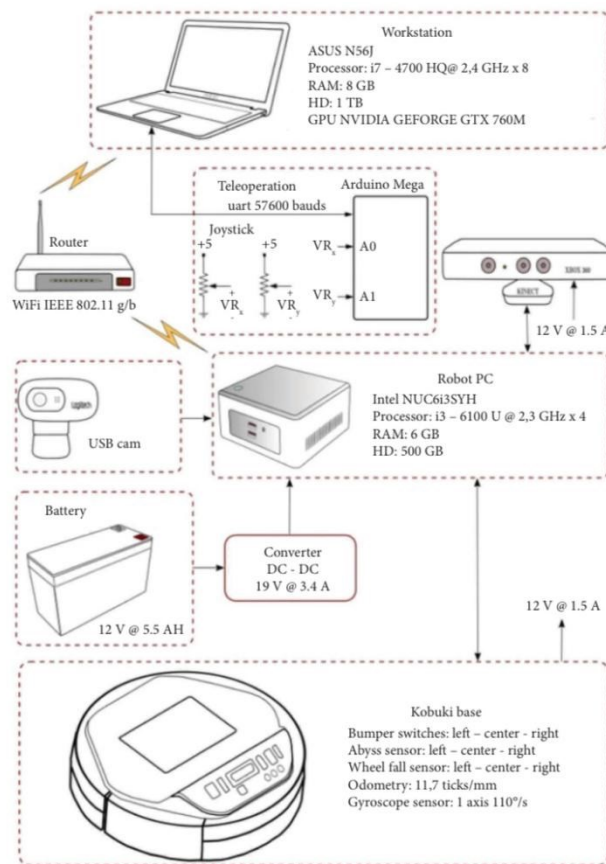


Figure 4. RobSLAM robotic system hardware architecture

4. Conclusion

This paper outlines the current development of logistics robots, and discusses the current limitations of logistics robots from the aspect of vision SLAM and laser SLAM. The development direction of multi-sensor fusion and optimized SLAM algorithm is proposed. Nowadays, the growth of e-commerce platforms, the demand for logistics robots also rose. Various SLAM methods are also widely used in robots nowadays, from the overall point of view, the application of single sensor slam tends to mature, and multi-sensor fusion slam will become the future development orientation of logistics robots. Multi-sensor fusion SLAM will also make the map construction and path planning of logistics robots more perfect. Future work could explore more effective route planning algorithms for more complex environments and more dynamic obstacles. Meanwhile, the continuous development of machine learning technologies is expected to lead to more innovative obstacle handling methods. Future research can be expanded to more practical application situations. The upcoming research will further advance the process of intelligent logistics robotics, unlocking the potential and creating opportunities for practical applications.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

- [1] Jia, G.; Li, X.; Zhang, D.; Xu, W.; Lv, H.; Shi, Y.; Cai, M. Visual-SLAM Classical Framework and Key Techniques: A Review. *Sensors* 2022, 22, 4582.
- [2] Smith, R.C.; Cheeseman, P. On the Representation and Estimation of Spatial Uncertainty. *Int. J. Robot. Res.* 1986, 5, 56–68.

- [3] Szendy, B.; Balázs, E.; Szabó-Resch, M.Z.; Vamossy, Z. Simultaneous Localization and Mapping with TurtleBotII. In Proceedings of the CINTI, the 16th IEEE International Symposium on IEEE, Budapest, Hungary, 19–21 November 2015.
- [4] Ul Islam, Q., Ibrahim, H., Chin, P.K., Lim, K. and Abdullah, M.Z. (2023), FADM-SLAM: a fast and accurate dynamic intelligent motion SLAM for autonomous robot exploration involving movable objects, *Robotic Intelligence and Automation*, Vol. 43 No. 3, pp. 254-266
- [5] I. Maurović, M. Seder, K. Lenac and I. Petrović, Path Planning for Active SLAM Based on the D* Algorithm With Negative Edge Weights, in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 8, pp. 1321-1331, Aug. 2018
- [6] Grisetti, G. , C. Stachniss , and W. Burgard . Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters. *IEEE Transactions on Robotics* 23(2007):p.34-46.
- [7] Zhang, Ji , and S. Singh . LOAM: Lidar Odometry and Mapping in Real-time. (2014).
- [8] Davison, Andrew J , et al. MonoSLAM: real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.6(2007):1052-1067.
- [9] Klein, Georg , and D. Murray . Parallel Tracking and Mapping for Small AR Workspaces." *IEEE & Acm International Symposium on Mixed & Augmented Reality ACM*, 2008.
- [10] Engel, Jakob , T. Schps , and D. Cremers . LSD-SLAM: Large-scale direct monocular SLAM." *European Conference on Computer Vision Springer*, Cham, 2014.
- [11] Mur-Artal, Raul , J. M. M. Montiel , and J. D. Tardos . ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Transactions on Robotics* 31.5(2015):1147-1163.
- [12] Mur-Artal, Raul , and J. D. Tardos . ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. (2016).
- [13] Mur-Artal, Raul , and J. D. Tardos . Visual-Inertial Monocular SLAM with Map Reuse. *IEEE Robotics and Automation Letters* PP.99(2016):796-803.
- [14] Tong, Qin , L. Peiliang , and S. Shaojie . VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE Transactions on Robotics* PP.99(2017):1-17.
- [15] Jiang G, Yin L, Jin S, Tian C, Ma X, Ou Y. A Simultaneous Localization and Mapping (SLAM) Framework for 2.5D Map Building Based on Low-Cost LiDAR and Vision Fusion. *Applied Sciences*. 2019; 9(10):2105.
- [16] Tourani A, Bavle H, Sanchez-Lopez JL, Voos H. Visual SLAM: What Are the Current Trends and What to Expect? *Sensors*. 2022; 22(23):9297.
- [17] Zou, Danping , and P. Tan . CoSLAM: collaborative visual SLAM in dynamic environments. *IEEE Transactions on Software Engineering* (2012).
- [18] Qu Y, Yang M, Zhang J, Xie W, Qiang B, Chen J. An Outline of Multi-Sensor Fusion Methods for Mobile Agents Indoor Navigation. *Sensors*. 2021; 21(5):1605.
- [19] Lou L, Li Y, Zhang Q, Wei H. SLAM and 3D Semantic Reconstruction Based on the Fusion of Lidar and Monocular Vision. *Sensors*. 2023; 23(3):1502.
- [20] Peng H, Zhao Z, Wang L. A Review of Dynamic Object Filtering in SLAM Based on 3D LiDAR. *Sensors*. 2024; 24(2):645.
- [21] Han Z (2023) Multimodal intelligent logistics robot combining 3D CNN, LSTM, and visual SLAM for path planning and control.*Front. Neurorobot.* 17:1285673.
- [22] Cheng, X. and C. Wang. 2024. A SLAM and Deep Vision Computing-based Robotic Path Planning Method for Agricultural Logistics Warehouse Management. *Pakistan Journal Agricultural Sciences* 61:769-779.
- [23] Gustavo A. Acosta-Amaya, Juan M. Cadavid-Jimenez, Jovani A. Jimenez-Builes. Three-Dimensional Location and Mapping Analysis in Mobile Robotics Based on Visual SLAM Methods. *Journal of Robotics* Volume 2023, Article ID 6630038, 15 pages.