Proceedings of the 4th International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/2026.33031

Application and Development of Image Recognition in the Navigation of Logistics Robots

Yiran Zheng

School of Mechanical Engineering, Shandong University, Jinan, China
202400161061@mail.sdu.edu.cn

Abstract. With the rapid expansion of e-commerce logistics, traditional fixed-route navigation methods used in logistics robots are increasingly constrained by low flexibility, high maintenance costs, and limited robustness in dynamic environments. To address these challenges, this paper reviews the application and development of image recognition technology in the navigation stage of logistics robots. The study first analyses the limitations of conventional navigation methods, including magnetic stripe, QR code, and track-guided navigation, and then examines the role of image recognition in key navigation stages such as localisation and mapping, path planning, and obstacle avoidance. On this basis, the paper summarises the iterative evolution of image recognition technologies from traditional image processing methods to deep learning approaches represented by convolutional neural networks and YOLO-based object detection. Their respective advantages in perception accuracy, environmental adaptability, and real-time performance are discussed. The paper further explores future development trends, including multi-sensor fusion, 5G-enabled communication, edge-cloud collaborative architectures, and path-planning optimisation. It concludes that image recognition has become a crucial technical foundation for improving the intelligence, autonomy, and operational efficiency of logistics robots, although challenges remain in complex lighting conditions, irregular object recognition, and environmental adaptability.

Keywords: Logistics Robots, Image Recognition, Autonomous Navigation

1. Introduction

According to the report released by the National Development and Reform Commission, the total social logistics volume of China reached 352.4 trillion yuan in 2023 and rose to 360.6 trillion yuan in 2024, with these statistics going up year by year. The continued growth of e-commerce and cross-regional fulfilment has steadily driven up demand for logistics services.

Fixed-route navigation methods that lack image recognition are widely used in traditional robots, which were initially deployed in simple warehousing environments. However, as the operational scenarios become increasingly complex, significant challenges have been revealed in this method. Therefore, image recognition technology, which is faster, better and cheaper, comes to a crucial solution to break through the difficulties faced by logistics robot navigation.

Most of the traditional fixed-route planning robots have their problems with navigation, which may be categorised under the efficiency, robustness, and complexity of system design. Regarding efficiency, it will not be efficient in changing the dynamic routes of cargo or distribution of obstacles, thus the robots will fail under congestion or due to changes in tasks. In the case of robustness, when the highly depended-upon preset markers get broken, they will lead to navigation issues, including positioning errors, and hence, ensuring that the robots navigate complex logistics will be challenging. On the issue of complexity in a system design, these markers, in terms of layout, installation and maintenance of their layouts, require a lot of manpower and materials, and on top of that, when the warehouse layouts are changed, there is a high cost involved in changing the routes, thus raising system design costs overall.

The above navigation issues can be the cause of the image recognition technology. Through the visual sensors, the robot can capture real-time scene imagery, and the technology will identify the location of the goods and obstacles, and then the technology will provide the important information that is used to plan optimally the navigation path, which makes the robot's steps more efficient. Moreover, it uses intrinsic scene features during localisation as opposed to routing, hence enhancing its flexibility to highly challenging environments. The new robot does not require fixed route markers, which further streamlines the system design process, also helps in cost reduction as a result of modifying a route and periodic maintenance, and consequently reduces the entry barrier to operating the system. The deep-seated combination of image recognition and robot navigation in the field of logistics may overcome significant disadvantages of traditional robots, thus leading to the intelligent carving out of logistics robots.

The paper will discuss the use of image recognition during the navigation stage of logistics robots and how it can be progressed to overcome traditional issues in navigation, and will also review a wide range of topics from three perspectives. With an aim of identifying the role of image recognition in the navigation of the logistic robots as well as analyzing how it has the potential to solve the traditional navigation-related challenges, this article will discuss and provide summaries of the present status of image recognition in the area as well as its future prospects, and present insights and prospects to readers to learn more on the subject.

2. Image recognition technology in logistics robots

2.1. Logistics robot specialising in the conventional image recognition technology

Magnetic Stripe Navigation Technology. The principle of magnetic stripe navigation is based on the fact that the magnetic field is initiated by magnetic stripes laid over the ground by creating a magnetic field on the bottom of the robot, which is sensed by the electromagnetic induction sensors, which translate the magnetic signals into navigation commands. The most fundamental concept of this technology is the recognition, as well as the decoding of electromagnetic induction signals [1]. Magnetic stripes are normally positioned following certain paths, and thus the robot will automatically know where it is and move in the correct direction by sensing magnetic cues in its magnetic sensor [2]. Nevertheless, the disadvantages of this technology are high maintenance and low flexibility.

QR Code Navigation Technology. QR code navigation involves the visual sensors that are installed on the robot and are used to scan the ground QR codes using the phones. Once the images have been decoded, it gets location coordinate information, thus making it able to position accurately [3]. Nonetheless, there are two significant disadvantages of this technology: inadequate environmental adaptability and unprofitability.

DOI: 10.54254/2755-2721/2026.33031

Track-guided Navigation Technology. Track-guided navigation limits the exploring path of a robot on the basis of physical tracks, and it is grounded on the use of sensors incorporated on the tracks to determine where a robot is located [4,5]. Track-based navigation allows movement of the equipment by laying tracks on the ground, but both these systems have significant weaknesses in that they occupy a lot of space and are not very flexible.

2.2. Navigation in logistics robots in stages

Localisation and Mapping Stage. The main technologies used in autonomous navigation are localisation and mapping, and consist mainly of SLAM algorithms and environmental modelling. The SLAM algorithms may be classified into the Gmapping and Car-tographer algorithms. The gmapping algorithm is developed using a particle-filtering SLAM system that separates localisation and mapping. It does very robust mapping in four steps that include sampling, importance weighting, resampling, and map estimation [6,7]. Comparatively, the Cartographer algorithm takes a graph-optimisation approach, where LiDAR pulling data is used to create local maps, and loop closure identification decreases the computing burden [8]. There are two broad steps in environmental modelling, namely the grid-based modelling and the topological-layer optimisation. The working environment grid-based modelling divides the work environment into a two-dimensional grid map, which is continually updated in real-time, with the aid of radar and visual sensors to produce an environmental model and identify regions of obstacles [4,5]. Following this, topology-layer optimisation is done, based on which a topological map is established on the grid map and the computational complexity is reduced by a node combining and eliminating redundant nodes process [6].

Path Planning Stage. Path planning aims at identifying the optimal route between the point of origin and the destination. The algorithm is the major component of this phase. Two major algorithms applied in this phase are A* and DWA. The A* algorithm is a type of search that relies on heuristic search, which utilises an evaluation function to estimate the path between the starting position and the target, and which is defined by the characteristics of a global optimality path [9]. Instead, the DWA algorithm integrates the dynamic window algorithm to maximise the local path and minimise the number of turns and enhance the smoothness of the path [10].

Obstacle Avoidance Stage. Robot obstacle-avoidance module majorly depends on three directions, namely sensor technology, navigation technology and deep learning and reinforcement learning. First, sensor technology: Sensor technology helps logistics robots gather the latest environmental data, through which they can plan a path and avoid obstacles [11]. Second, navigation technology: Autonomous navigation and obstacle avoidance robots utilise technologies, which include LiDAR and SLAM algorithms [12]. Third, deep learning and reinforcement learning: Deep neural networks are used in the case of dynamic obstacle avoidance [13].

3. Image recognition technology specification differences with the process of iteration

Image processing and recognition technologies are the most important in image recognition in the navigation phase of the logistics robots.

3.1. Traditional methods

The traditional image recognition processes usually have three main steps that are followed by one being the preprocessing, feature extraction and the last being classification and recognition.

DOI: 10.54254/2755-2721/2026.33031

Some of the techniques commonly employed during the preprocessing stage are converting to grayscale, noise reduction filtering, and illumination correction. Grayscale conversion is used to simplify computational complexity since an RGB image is converted into a grayscale image [14-17]. The filtering used to reduce noise works well with median and Gaussian filtering at removing salt-and-pepper noise and Gaussian noise, respectively, thus removing denoising [18]. The Gamma correction through illumination correction helps to alleviate light imbalance conditions to increase contrast [19]. The discriminative information in an image is obtained in feature extraction. With reference to the type of features, there are three types of features, namely local, global, and holistic. Some of the algorithms in use are SIFT, HOG and LBP. Classification and recognition are normally obtained through the employment of classifiers that make decisions based on extracted features.

The conventional techniques cannot go on in the absence of the human touch. Thus, this technology can be used in the logistics sector only in cases that have low computational needs and fairly simple processes. With the assistance of infrared sensors or cameras, robots are able to take ground pictures, identify guiding lines, and ascertain their position and direction in which they are moving, depending on the particular marks painted on the ground, thus producing robot motion.

3.2. Deep learning methods

Deep learning methods are more robust and recognise accurately in comparison with traditional methods. This renders deep learning strategies applicable in different and challenging environments faced by logistics robots.

First, the sorting of the logistics and the identification of the packages. Deep learning in automated sorting systems is applied in object category recognition and three-dimensional localisation. As an illustration, a deep-learning-based automatic recognition and sorting robot trains texture features of coal and gangue with the help of visual perception, and it coordinates the actions of various robotic arms to reach an accurate sorting [14].

Second, detection of obstacles and path planning. The logistics robots should be able to evade real-time obstacles in a dynamic environment. With the YOLO algorithm used as an object detection algorithm, an edge detector is capable of detecting edge information of the obstacles rapidly, and finding an effective route by using data obtained by the pose sensor [15].

Third, multi-sensor fusion and edge computing. Edge machine vision cameras decrease the computational cost by using lightweight algorithms, which is beneficial in providing real-time decisions to the robots [16].

Employs a series of convolutional filters to extract characteristics, producing an output map that matches the input map in appearance and represents object features within the image.

Convolutional Neural Networks (CNNs) are an algorithm that takes advantage of filtering convolutions to generate features that give an output map that looks like the input map and captures object features in the image.

The Convolutional Neural Network (CNN), as a prime example of deep learning, allows achieving a much lower level of network complexity, the number of network parameters, and greater resistance to transformation, such as translation, scaling, etc., due to the following factors: the use of local and point-to-point connections, weight sharing, pooling, and the multi-layer network. Its basic building block is convolutional layers (local features are removed with the help of sliding convolutional kernels), pooling layers (max-pooling and dimensional reduction), and fully connected layers (global information integration and classification) [20]. These networks learn local image features based on convolutional layers with high recognition accuracy in straightforward sorting. Nevertheless, the shortcomings of convolutional operators prevent them from sufficiently

DOI: 10.54254/2755-2721/2026.33031

capturing long-range contextual information in spatio-temporal images; as a result, shallow layers are unable to acquire effective features [21].

You Only Look Once (YOLO). The main concept of YOLO assumes that the object detection process can be presented as a single regression, directly regressing the validation of the stage of predicting a bounding box and class probabilities on the entire image.

In comparison to the conventional object detection algorithms, the YOLO algorithm attains end-to-end real-time detection and enjoys an outstanding balance between speed and accuracy. Prior to YOLO, the majority of object detection algorithms were based on sliding windows or region proposal techniques, where one needed to search through various places on a photograph and was, therefore, computationally intensive.

The workflow of the YOLO algorithm can be broken down into three essential tasks: it breaks the input image into an S×S grid, then each cell is expected to predict B bounding boxes and the corresponding classification ratings, and non-maximum suppression (NMS) is applied to get the best possible boxes of both predictions, which results in final detection outputs. Such a single-detect architecture allows YOLO to perceive the images in a more contextual, global way that greatly reduces the chances of confusion between the background and the objects, having a huge processing speed of over 45 frames per second that forms the basis of real-time logistics implementation [22].

The fundamental technological breakthrough of YOLO is the deep convolutional neural network structure, which is especially suitable for logistics robots. The base network uses 24 convolutional layers to get features and then takes two fully connected layers to predict. This greatly minimises redundancy in computation in the detection process. YOLO has a distinct benefit over RCNN and SDD algorithms; it can attain a faster convergence speed, meaning it can detect faster, in addition, it can be more accurate on small stacked objects like courier boxes [22]. All of these features precondition the uniqueness of the YOLO family of algorithms in the context of automation in logistics, particularly in time-sensitive process automation, e.g., package sorting and cargo identification.

4. Trends in the future development of the navigation segment

4.1. Multi -sensor fusion technology

Logistics robots have the ability to move towards multi-sensor fusion technologies as opposed to single sensors. The coordinated implementation of devices like LiDAR, visual sensors, and magnetic sensors can contribute to the enhancement of navigation accuracy and obstacle-avoidance possibilities. The use of laser SLAM-based navigation technology facilitates real-time path planning in complicated conditions, whereas the use of visual SLAM technology adds extra flexibility to the ability of robots in dynamic environments [23].

Furthermore, the 5G technology, with its advantages of a huge bandwidth, a high level of connectivity, and minimal latency, has totally reformed the communication framework of the logistics robots. Firms such as Geek+ have already released 5 G-based AMR operating systems, which permit scheduling of a multi-robot cluster, and real-time collaborative control [24].

4.2. Path planning optimisation

Crossover between A* and DWA algorithms is a mainstream method of path planning. The two algorithms can also be improved. A* algorithm has the ability to traverse the turning nodes on the original path, and when there are no obstacles, the current turning node can be replaced with a

DOI: 10.54254/2755-2721/2026.33031

diagonal node, thus the number of turns decreases as well as the cumulative turning angle in the local path [25]. Meanwhile, a turn penalty term may also be used in the A* algorithm, but dynamically the heuristic function is used to reduce the length of the path-searching [25]. DWA algorithm is able to add a constraint to its evaluation function in order to avoid drastic path changes and improve its dynamic obstacle-avoidance performance [26].

4.3. Technology integration between domains

The collaboration structure is edge-based, which takes advantage of the clear cut between edge machine vision and cloud-based AI models in order to realise real-time data processing and quick response. Lightweight algorithms of edge devices process image data, which is in turn uploaded to the cloud to undergo intensive analysis, hence changing complex image data into simple decisions and directives. This also maximises both motion and trajectories, minimising the number of failures [27].

4.4. Challenges

Presently, image recognition can be used in navigation, but it does not deal with complex objects; this is a bottleneck, considering that its use in this field still requires a lack of accuracy in the algorithm, and there is a lack of environmental adaptation features. In particular, it is susceptible to errors in various conditions of changing light, distortion of objects, or interference of shadows, and it is unable to deal with irregular objects like liquids, bulky goods, or even delicate objects [28].

5. Conclusion

Image recognition has a very significant place in the logistics robots' navigation stage. To overcome the limitations of the conventional system of magnetic stripe and QR-based navigation, there is image recognition based on deep learning, which successfully overcomes the obstacles on the path, traffic and planning of vectors: due to its strong environmental perception and the benefits of real-time processing, the problem of recognising objects is tackled, and the reactions to dynamic obstacles are avoided. The intensive multi-sensor fusion, 5G communication technologies, and edge collaborative architecture approaches will allow the future logistics robots to have greater environmental flexibility and ability to collaborate in a cluster, as such, despite the limitations of modern technologies in response to extreme lighting conditions and objects with irregular shapes. Thus, the constant enhancement of the innovative use of image recognition technology during the navigation stage is not only one of the primary avenues of increasing the level of automation in logistics but also an unavoidable option in the development of smart logistics of high quality.

References

[1] Yang, G., Wei, J. (2022). Path planning of logistics robot based on improved A* and DWA algorithm. Science Technology and Engineering 22(34), 15213–15220
[2] Yu, H. (2021). Design of automatic transportation and transfer unit based on AGV. Mechanical Management and Development 36(1), 187–189
[3] Zhang, H., Zhang, C., Yang, W., Chen, C. (2015). Localization and navigation using QR code for mobile robot in indoor environment. In: Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 2501–2506
[4] Elfes, A. (1987). Sonar-based real-world mapping and navigation. IEEE Journal on Robotics and Automation 3(3), 249–265

DOI: 10.54254/2755-2721/2026.33031

[5] Konolige, K. (1997). Improved occupancy grids for map building. Autonomous Robots 4(4), 351–367
[6] Choset, H., Nagatani, K. (2001). Topological simultaneous localization and mapping (SLAM): Toward exact localization without explicit localization. IEEE Transactions on Robotics and Automation 17(2), 125–137
[7] Grisetti, G., Stachniss, C., Burgard, W. (2007). Improved techniques for grid mapping with Rao–Blackwellized particle filters. IEEE Transactions on Robotics 23(1), 34–46
[8] Hess, W., Kohler, D., Rapp, H., Andor, D. (2016). Real-time loop closure in 2D LIDAR SLAM. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 1271–1278
[9] Hart, P., Nilsson, N., Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. IEEE Transactions on Systems Science and Cybernetics 4(2), 100–107
[10] Fox, D., Burgard, W., Thrun, S. (1997). The dynamic window approach to collision avoidance. IEEE Robotics and Automation Magazine 4(1), 23–33
[11] Ušinskis, V., Nowicki, M., Dzedzickis, A., Bučinskas, V. (2025). Sensor-fusion-based navigation for an autonomous mobile robot. Sensors 25(4), 1248
[12] Yue, X., Zhang, Y., Chen, J., Chen, J., Zhou, X., He, M. (2024). LiDAR-based SLAM for robotic mapping: State of the art and new frontiers. Industrial Robot: The International Journal of Robotics Research and Application 51(2), 196–205
[13] Fan, J., Liu, Y. (2020). Design of a small solar-powered warehouse logistics robot. Scientific and Technological Innovation (25), 90–91
[14] Hu, Y., Li, Y., Zheng, X., Liu, J. (2022). Design of an automatic identification and sorting robot based on deep learning. Computer Programming Skills and Maintenance (6), 114–116
[15] Wu, Y., Liu, C., Xie, S., Ban, Y., Sui, X., Huang, Y., Zhang, Y. (2021). Robot environment perception and autonomous obstacle avoidance based on visual deep learning. Electronic Measurement Technology 44(20), 99–106
[16] Li, H., Du, K., Xing, Z., Wang, G., Zhang, H., Liu, G. (2024). Precise localization of logistics AGV based on multi-sensor fusion and its application in human–machine safety. Manufacturing Technology and Machine Tool (5), 63–69
[17] Wu, Y. (2017). Development and application of image recognition technology. Electronic Technology and Software Engineering (4), 86
[18] Huang, Z., Li, J. (2020). Paper of image recognition research based on convolutional neural networks. Automotive Engineer (10), 11–13
[19] Yang, J., Shi, Y., Xiong, X. (2006). Improved gamma correction method for illumination processing. Journal of Civil Aviation University of China (4), 39–42
[20] LeCun, Y., Bengio, Y., Hinton, G. (2015). Deep learning. Nature 521, 436–444
[21] Liu, R., Lehman, J., Molino, P., Petroski Such, F., Frank, E., Sergeev, A., Yosinski, J. (2018). An intriguing failing of convolutional neural networks and the CoordConv solution. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 31
[22] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). this paper only looks once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788
[23] Wang, L., Liu, K., Yao, Y., Zhou, W. (2024). Autonomous navigation of logistics sorting robot based on laser SLAM. Laser Journal 45(5), 236–240
[24] Hampel, G., Li, C., Li, J. (2019). 5G ultra-reliable low-latency communications in factory automation leveraging licensed and unlicensed bands. IEEE Communications Magazine 57(5), 117–123
[25] Song, Y., Zhang, H., Cheng, C. (2023). Research on logistics robot algorithm based on improved A* algorithm. Journal of Changchun University of Technology 44(2), 189–192
[26] Qin, H., Shao, S., Wang, T., Li, Y., Wang, N., Yao, C. (2022). An improved dynamic window approach for mobile robot dynamic path planning. In: Proceedings of the International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), pp. 348–353
[27] Zhu, Y., Ling, Z., Zhang, Y. (2020). Research progress and prospects of machine vision technology. Journal of Graphics 41(6), 871–890
[28] Mao, J., Li, T., He, Y., Guo, Y. (2024). Application prospects of machine vision recognition robots in the packaging sorting field. China Packaging 44(6), 34–38