Review on Autonomous Robot Mobility Based on Visual Deep Learning

Xun He[1], Tianyi Xu[2] and Lo San Yuan[3,*]

[1] College of Engineering，HuaZhong Agricultural University, Wuhan,430070,China
[2] Mechanical Engineering and Automation，Hubei Polytechnic University，Huangshi，435003, China
[3] Faculty of Engineering, The University of Hong Kong, Hong Kong, China
*losan.yuan@outlook.com

Abstract. The topic related to dynamic autonomous navigation in complex environments has become more popular in the robotics research area. It is very crucial for mobile robotics to have automatic obstacle avoidance and path planning. Traditional methods, such as simultaneous localization and mapping (SLAM) technology, for instance, typically rely on sensors such as cameras, LiDAR, and ultrasonic sensors, in order to accomplish local obstacle avoidance. While these approaches demonstrate high reliability, they may be constrained by limitations in flexibility, detection efficiency, and real-time performance. In recent years, numerous emerging technologies about artificial intelligence (AI) and deep learning are introduced and applied in related aspects. By combining those state-of-the-art techniques and machine vision, the capabilities of perception and motion decision making for robotics are dramatically enhanced. In this article, the application of deep learning models, such as convolutional neural networks, to the problem of path planning, including obstacle avoidance algorithms of mobile robots will be mainly discussed. Key issues and challenges in current research and potential solutions along with future development trends will be determined at the end.

Keywords: Autonomous robot movement; deep visual learning; automatic obstacle avoidance; path optimization.

1 Introduction

With the development of the times, robots are more and more important nowadays. For example, in the industrial field, robots can avoid obstacles rapidly with autonomous movement, so that accidents can be avoided. In some service industries like hotels and restaurants, this robot can arrive at the designated location and finish the task precisely. This has brought convenience to humans.

In recent years, the field of robotics has developed rapidly. Visual deep learning technology dominates the technological market. For example, Tencent developed a Quadruped mobile robot named Jamoca, which can walk, jump, trot and perform automatic obstacle avoidance [1]. West Lake University developed an insent-scale soft

K. Subramanian (ed.), *Proceedings of the International Workshop on Advances in Deep Learning for Image Analysis and Computer Vision (IWADIC 2025)*, Advances in Computer Science Research 128,
https://doi.org/10.2991/978-94-6239-648-7_35

robot which can still move after falling from a height of 108 metres. Besides, it can jump continuously in the grass after swimming by itself for an hour [2].

This article presents research present Visual Deep Learning applied in robots moving autonomously. Moving autonomously includes the robot's automatic obstacle avoidance technology when it faces an obstruction and the robot's path optimization technology that can choose the best path.

2 Move Autonomously

According to the process of robots with visual deep learning moving autonomously, we can divide it into visual perception, decision-making, planning and executive feedback. Visual perception is when robots capture and collect information about the surrounding environment through a camera, and then examine surrounding obstructions. But it may reduce work efficiency because of its slow speed and long time. Executive feedback is that robots drive their power system in order to move through visual perception according to the planned path. But when confronted with emergencies,it may not perceive and execute feedback rapidly. Using decision-making planning can make robots conduct visual deep learning through the result of visual perception, thereby prejudging the dynamic movement of obstruction, providing executive feedback in advance, improving the work efficiency of robots and ensuring the safety during the robots' automatic movement. These are the important development directions of the visual deep learning robot.

In the automatic movement of a visual deep learning robot, automatic obstacle avoidance and path optimization technology are the key issues. Foreign scientific researchers establish a neural network model for obstacle detection and enhance the learning algorithm(Q-learning) through analyzing the process of robots' automatic movement in different environmental conditions. It allows robots to have the ability for visual deep learning, and robots can respond quickly and avoid obstacles when they face an obstruction. Thereby making the path optimization.

3 Obstacle Avoidance

3.1 Object Detection

Object detection is a core task in computer vision. To automatically avoid obstacles, mobile robots must detect surrounding obstacles and obtain accurate information, such as shape, distance, and specific coordinate position. However, traditional detection methods are easily affected by external factors such as environmental complexity or lighting changes, which can lead to a decrease in detection accuracy and affect obstacle avoidance.

YOLO for Recognition and Detection. For the purpose of improving the accuracy and efficiency of robot obstacle detection, convolutional neural networks (CNNs) have become the primary technology for computer vision processing tasks. In 2015, Joseph

Redmon proposed the You Only Look Once (YOLO) model, which aims to achieve one-stage, end-to-end real-time object detection using a single convolutional neural network model. The YOLO model balances the accuracy and efficiency of robot obstacle detection. After years of iteration, as shown in Table 1 [3], different versions have their own unique performance parameters.

Table 1. Features of each version of the YOLO model [3]

Version	Characteristics
YOLOv2	Batch normalization (BN) was introduced to improve training accuracy A high-resolution classifier was used Model: Darknet-19 (containing 5 max pooling layers and 19 convolutional layers)
YOLOv3	Multi-label classification using logistic regression DBL module combines convolutional layers, activation functions, and batch normalization Model: Darknet-53 (containing 53 convolutional layers)
YOLOv5	Similar to YOLOv4's internal structure Significantly improved training speed and accuracy, and easier to use
YOLOv7	Outperforms all known detectors in the 5–160 FPS range Reduces training time while maintaining accuracy and performance
YOLOv8	Supports multiple vision tasks with low training cost Includes five detection modules and one prediction layer Easy to deploy Surpasses previous models in accuracy and efficiency

Application of YOLO. There are several versions of detection models used in the field of machine vision. In order to balance the computing power of robot hardware equipment and model efficiency, Liu Suyi proposed a neural network pruning algorithm for the YOLOv3 algorithm, which is to simplify and compress the existing target detection algorithm network [4]. After retraining, compared with the original YOLOv3 model, the size of the compressed detection model and the memory occupied during run-time were reduced by about 40% and 35% respectively, the detection speed was increased by about 30%, and the detection accuracy remained the same. Lan Renwen used the YOLOv5s detection model to realize the robot's obstacle visual perception. After training the model using PyTorch on an NVIDIA GeForce RTX 2080Ti GPU, its accuracy reached 94% [5]. Seema Duhan and Ruchi Panwar used the YOLOv8 model to identify small obstacles in the hospital environment to ensure the safe movement of the robot [1]. The YOLOv8 has a detection accuracy of 92.5% in complex hospital environments, which is an average improvement of 12% over previous generations. In addition, the model can maintain high detection accuracy in low-light conditions at night [3]. In general, the YOLO target detection method has the characteristics of fast speed, high accuracy, and ease of use. Therefore, this method has been widely used and studied in robot vision. In addition, due to YOLO's open source nature, academia as well as a variety of industries have actively contributed to its development and application, which has led to advancements in a number of areas such as unmanned driving, drones, and automation.

3.2 In-depth intensive learning

In-depth intensive learning is combined deep learning and intensive learning. In the process of robots' automatic movement,the robots obtain data of this situation through camera, radar and other sensors.And then robots can extract the data characteristics of its environment through the Cvolutional Neural Network and other deep learning model. Finally, the robots can change its driving route facing changeable dynamic environment combined with intensive learning to achieve automatic obstacle avoidance.

Q-leaning. Q-learning proceed to the next step through judge the value of each step. In the process of robots' automatic movement,this algorithm will make a Q data form and take notes of the robots' driving status by their moving constantly, from which the best location is chosen(the highest Q data). And then a rewarding mechanism is set up. After this, the Q data from is renewed according to the new state during the robots' driving and the rewards obtained by robots.

Different from other algorithms,Q-learning doesn't require model and can change robots' movement status only by calculating the Q data. And the learning strategy and execution strategy of this algorithm are separated,s o it can change strategies to different situations and the data collection and utilization are more efficient.

Q-learning in the decision-making process in the Maldives or in the SMP can be used in the quantitative investment [6]. Solve the scheduling problem of the workshop in the dynamic environment. Using Q-learning algorithm can improve performance indicators [7]. Using Q-learning algorithms can adapt through adjusting the Q data not only in the continuous state but also the discrete state.And its code flow is clear,which can build model quickly.

Deep Q Network. Deep Q Network(DQN) is combine Cvolutional neural network and classical Q-learning algorithms.Compared with classical Q-learning algorithms l,Deep Q Network is more convenient and practical. Classical Q-learning mainly relies on the Q data form which formed by the state shown in the course of exercise in the apply about the automatic obstacle avoidance,but when it faces some continuous and complex motion states,Q data form cannot calculate efficiently because it has recorded too much data. More importantly,in some very noisy environment,classical Q-learning algorithms will be disturbed and its data collection is unstable, efficiency of work is low. And when dealing with layout for some obstacles that have never been encountered, classical Q-learning algorithms needs relearn from the beginning and it will take a long time.

For noise processing, DQN algorithms has a buffer,it takes the collected data storing in the buffer. And then it can reduce the impact of noise on data at any time. DQN has neural network, so it has generalization ability. For the distribution of obstacles with a certain degree of similarity,it can get reasonable Q data.DQN doesn't need to relearn when it faces unfamiliar obstacles.

In order to improve the operating efficiency of aircraft in the taxiway,using the DQN algorithm, two independent intelligent bodies named agent were studied to perform intelligent speed regulation of aircraft applying to use the same taxiway intersection at the intersection of the taxiway at the airport scene [8]. Nowadays,the prospects of DQN algorithms are good but it still has disadvantages. Because at the beginning of the test,DQN algorithms are not familiar with a certain environment,so it needs a lot of familiarization and data accumulation of the environment. As a result, these will take high cost of error and a lot of time and energy at the beginning of the project.

4 Technologies in Path Optimization

During the execution of navigation tasks, robots are typically equipped with a variety of intelligent devices. Relying on navigation technologies, these devices enable a series of key operations: modeling the surrounding environment, accurately locating the robot's own position, controlling its movement state, and detecting and avoiding obstacles. Among these operations, the core of safe path planning lies in effectively detecting and avoiding obstacles—and this is also a core function that any navigation technology must possess.

4.1 A* algorithm

T* algorithm was proposed by Khalidi, which combines the A* with path planning based on LTL.When solving LTL path planning problems in large 2D and 3D workspaces,this algorithm decreases both the quantity of nodes as well as the time taken for path generation in comparison with conventional algorithms [9]. Zhong Xunyu simplified the distance cost calculation as part of the risk cost function, brought down the number of nodes, and extracted the key path points.This method not only enable to avoid obstale but also can meet the real-time requirements of moving robots operating in large-scale environments [10]. Li Ye put forward a navigation algorithm which is used in deep sea.This approach draws on terrain information for a dynamic matching method to lower time consumption [11]. Sang Hongqiang's method avoids detours and frequent turns of mobile robots, reduces the count of search points near obstacles to ensure a safe distance, and identifies redundant path points by checking for surrounding obstacles—thereby optimizing the generated trajectory. This approach shortens the path length, reduces search time, and decreases the number of nodes [12].

4.2 Ant Colony Optimization (ACO)

Fang Shengkun et al. proposed an optimization strategy for ACO. This strategy improves the performance of the ACO algorithm by updating random pheromones and increasing the initial states of cycles,and the revised algorithm sees more effective use in automatic path planning for hole groups in machining [13]. Cui Junguo et al. introduced MsAACO to relieve the problems of insufficient convergence and inefficiency inherent in ACO. This approach enhances node selection performance, enabling efficient generation of smoother optimal path planning solutions. Meanwhile, the shorter path length and reduced turning time have improved the convergence efficiency and stability of ACO [14]. Liang Chuandong and colleagues put forward a IWOA-ACO,which can reduce operating loss of agricultural machinery and prolong driving mileage [15]. Li Guangxin proposed an IACO - IABC, which has advantages in reducing turns and path length, and improving the convergence speed of the algorithm [16].

4.3 Deep Learning Algorithms

Syed addressed the issue that most path planning algorithms fail to deal with by proposing a GAPCNN. By introducing directional autowave control based on dynamic threshold technology and accelerated neuronal firing, this network significantly reduces path query time [17]. Chu Zhenzhong proposed a DRL path planning method based on the DDQN to address the path planning issue of AUVs under the disturbance of ocean currents. A dynamic composite reward function was developed to enable AUVs to avoid obstacles and reach destinations, while enhancing the algorithm's accuracy and efficiency [18]. Zhou Yuting proposed a DRL-based autonomous robot path planning method for indoor blind spots to address the issue of blind spots in indoor environment exploration. This method realizes the exploration of blind spots while also enhancing the speed of convergence [19].

5 Conclusion

Automation and intelligence are the inevitable trend in robotic autonomous mobility development, with visual deep learning-driven autonomous mobility as its core cornerstone. The integration of YOLO-based visual models, deep reinforcement learning algorithms, and traditional path planning methods provides a promising technical approach for robots to perceive environments and make decisions. However, current technologies have shortcomings: they work in conventional scenarios but lack perception accuracy, decision precision, and execution coordination in extreme environments; the conflict between algorithm efficiency and hardware adaptability also fails to meet real-time and lightweight needs, limiting coverage of diverse scenarios.

With deep integration of deep learning and machine algorithms, improving algorithm efficiency is key for future progress. While robotic autonomous mobility now enables dynamic obstacle avoidance and path planning in some scenarios, it is far from generalized and intelligent. Urgent challenges include deepening multimodal fusion to enhance system generalization in unknown environments, and accelerating the "perception accuracy - decision efficiency - execution feedback" coordination to help robots adapt from specific scenarios to flexible real-world situations.

Authors Contribution

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

References

1. Tencent Announces Two New Advancements in Mobile Robots. Robot Technology and Application 2020(6), 10 (2020)
2. Jin, P.: Where's the "Machine Cockroaches". Guangming Daily, 2025-08-20 (007) (2025)
3. Duhan, S., Panwar, R.: YOLO Based Vision-Aided Obstacles Navigation and Avoidance with Path Planning Using SARSA Algorithm for Biped Robot in an Uncertain Hospital Environment. Indian Journal of Science and Technology 17(3), 236–246 (2024)
4. Liu, S.: Research on Obstacle Avoidance Method for Mobile Robots Based on Vision. Northeastern University (2020)
5. Lan, R.: Research on Robot Obstacle Avoidance Algorithm Based on Machine Vision. Automation Applications 65(13), 32–34 (2024)
6. Zhou, Q.: Q-learning Based on Semi-Markov Process and Its Application in Quantitative Investment. Zhengzhou University (2022)
7. Zhao, Y., Wang, Y., Zhang, J., et al.: Q Learning Algorithm in the Application of Job Shop Scheduling Problem. Journal of System Simulation 2022(6), 1247–1258 (2022)
8. Jiang, B., Li, Y., Liu, W., et al.: Cooperative Intelligent Speed Regulation Method for Airport Crossing Based on Deep Q-Learning. Aeronautical Computing Technology 52(1), 26–30 (2022)
9. Khalidi, D., Gujarathi, D., Saha, I.T.: A Heuristic Search Based Path Planning Algorithm for Temporal Logic Specifications. In: Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 8476–8482. IEEE, Paris (2020)
10. Zhong, X., Tian, J., Hu, H., Peng, X.: Hybrid Path Planning Based on Safe A* Algorithm and Adaptive Window Approach for Mobile Robot in Large-Scale Dynamic Environment. Journal of Intelligent Robotic Systems 99, 65–77 (2020)
11. Li, Y., Ma, T., Chen, P., Jiang, Y., Wang, R., Zhang, Q.: Autonomous Underwater Vehicle Optimal Path Planning Method for Seabed Terrain Matching Navigation. Ocean Engineering 133, 107–115 (2017)
12. Sang, H., You, Y., Sun, X., Zhou, Y., Liu, F.: The Hybrid Path Planning Algorithm Based on Improved A* and Artificial Potential Field for Unmanned Surface Vehicle Formations. Ocean Engineering 223, 108709 (2021)
13. Fang, S., Deng, Z., Long, D., et al.: Improved Strategy of Ant Colony Optimization for Path Planning via Stochastic Pheromone Updating and Cyclic Initialization. Journal of Mechanical Science and Technology 39(4), 2051–2062 (2025)
14. Cui, J., Wu, L., Xiao, W., et al.: Multi-strategy Adaptable Ant Colony Optimization Algorithm and Its Application in Robot Path Planning. Knowledge-Based Systems 288 (2024)
15. Liang, C., Pan, K., Lu, M., et al.: Multi-Node Path Planning of Electric Tractor Based on Improved Whale Optimization Algorithm and Ant Colony Algorithm (2023)
16. Li, G., Liu, C., Xiao, W., et al.: A Mixing Algorithm of ACO and ABC for Solving Path Planning of Mobile Robot. Applied Soft Computing 148 (2023)
17. Syed, U.A., Kunwar, F., Iqbal, M.: Guided Autowave Pulse Coupled Neural Network (GAPCNN) Based Real Time Path Planning and an Obstacle Avoidance Scheme for Mobile Robots. Robotics and Autonomous Systems 62, 474–486 (2014)

18. Chu, Z., Wang, F., Luo, C., et al.: Path Planning Based on Deep Reinforcement Learning for Autonomous Underwater Vehicles Under Ocean Current Disturbance. IEEE Transactions on Intelligent Vehicles 8(1), 108–120 (2023)
19. Zhou, Y., Yang, J., Lin, J.C.W., et al.: An Indoor Blind Area-Oriented Autonomous Robotic Path Planning Approach Using Deep Reinforcement Learning. Expert Systems with Applications 254 (2024)