Visual Navigation of Mobile Robots in Complex Environments Based on Distributed Deep Reinforcement Learning

Yi Zhang
Shenzhen Institute of Advanced Technology
Chinese Academy of Sciences
Shenzhen, China
y.zhang@siat.ac.cn

Wei Feng
Shenzhen Institute of Advanced Technology
Chinese Academy of Sciences
Shenzhen, China
wei.feng@siat.ac.cn

Zhile Yang
Shenzhen Institute of Advanced Technology
Chinese Academy of Sciences
Shenzhen, China
zl.yang@siat.ac.cn

Zhaokun Zhou
Shenzhen Institute of Advanced Technology
Chinese Academy of Sciences
Shenzhen, China
zk.zhou@siat.ac.cn

Zihan Zhu
Shenzhen Institute of Advanced Technology
Chinese Academy of Sciences
Shenzhen, China
zh.zhu@siat.ac.cn

Weijun Wang*
Shenzhen Institute of Advanced Technology
Chinese Academy of Sciences
Shenzhen, China
wj.wang@giat.ac.cn

2022 6th Asian Conference on Artificial Intelligence Technology (ACAIT) | 978-1-6654-5311-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ACAIT56212.2022.10137974

***Abstract*—The increasingly popular method of deep reinforcement learning can not only help mobile robots output accurate actions in complex environments but can also search for collision-free paths. In this paper, a robot visual navigation model in complex environments based on distributed deep reinforcement learning is proposed. According to the characteristics of different regions in the complex environment, the environment is divided into several regions, and we proposed method can realize visual navigation in large scene complex environments. In these regions, we combine long-short term memory (LSTM) and proximal policy optimization (PPO) algorithms as a local visual navigation model and design a new reward function that trains the target through factors such as the action of mobile robots, the distance between robots and the target, and the running time of robots. We create respective experience pool independently through model training. The model of robot visual navigation via distributed deep reinforcement learning uses the RGB-D image obtained from the first perspective of mobile robots and the polar coordinates of the target in mobile robots coordinate system as input, and the continuous motion of mobile robots as output to realize the task of end-to-end visual navigation without maps. Our model can complete accurately robot visual navigation in large complex scenes without maps and human intervention. In our experiments, we verify our proposed model by performing the promising navigation tasks in virtual environments.**

***Index Terms*—deep reinforcement learning; mobile robot; visual navigation**

*Corresponding author

978-1-6654-5311-0/22/$31.00 ©2022 IEEE

I. Introduction

With the continuous development of mobile robot navigation technologies, mobile robots are becoming more and more common both in the office and at home [1]. Among the numerous applications in smart vehicles and robotic systems, autonomous navigation technology is receiving increasing attention [2]. Most of the traditional mobile robot navigation is a map-based method, including the Simultaneous localization and mapping (SLAM) [3] and path planning [4]. Generally, the map of the environment is built by SLAM methods, and then according to the map a collision-free path from the starting point to the destination is generated by path planning methods [5]. In unknown, unstructured, and dynamic environments, SLAM-based methods become ineffective, and subsequent path planning is out of the question [6]. So mobile robots should have the ability to learn from the scene and make judgement to navigate to the set target position safely. With the rise of deep reinforcement learning (DRL), research on autonomous navigation of mobile robots in unknown, unstructured, and dynamic environments has attracted interest [7]. Due to the high cost of laser sensors, the autonomous navigation of mobile robot based on DRL is mostly based on visual sensor [8]. In DRL based visual navigation, the robot can interact with the moving and non-moving objects in the environment through visual information to gradually learn and then optimize the performance of its navigation

[9]. At present, visual navigation methods based on DRL are mostly targeted at indoor navigation, social navigation, local obstacles, and multi-robot navigation [10], and there are fewer research on indoor and outdoor navigation in complex scenes. In the previous algorithms of deep reinforcement learning, the large and complex environment will lead to the difficulty and slow convergence of deep reinforcement learning [11]. In addition, the more complex the navigation task, the algorithm will spend more time training.

In this paper, we discuss the problem of deep reinforcement learning based visual navigation and propose a new learning architecture that enables mobile robots to find a collision free path to the destination in the large complex scenes. The large-scale complex environment will lead to the explosive growth of model training time. In order to reduce the training time, we put the complex navigation task according to the region was divided into several simple tasks. In these regions, we combine long-short term memory (LSTM) and proximal policy optimization (PPO) algorithms as a local visual navigation model, and a new reward function is designed to train the target through factors such as the action of mobile robots, the distance between robots and the target, and the running time of robots. Agents are trained by the proposed robot visual navigation model. And then, the trained model is used to lead the mobile robot to complete the complex global navigation task, and complete complex navigation tasks by crossing multiple local areas through channels in complex environments. By the proposed method, the mobile robot can realize visual navigation in the large complex scene.

II. Related works

A. Deep Reinforcement Learning

DRL is considered one of the most promising approaches to artificial general intelligence. DRL has developed rapidly in recent years, an example of whose successful application is the Deep Q-Network (DQN) used in Atari games [12]. The current DRL algorithm is mainly divided into two categories: value-based and policy-based. The value-based DRL algorithm is represented by DQN. In order to solve the over estimation problem of target Q-network in DQN, the Double DQN (DDQN) algorithm is proposed, which uses different networks to calculate the value of target Q-network and select action respectively [13]. The dueling DQN algorithm is proposed by decomposing the Q-network into two networks which are the state-dependent action advantage function and the state value function [14]. DRL based policy function mainly includes Asynchronous Advantage Actor Critic (A3C) [15], deep deterministic policy gradient (DDPG) [16], Trust Region Policy Optimization (TRPO) [17], Proximal Policy Optimization (PPO) [18] and other methods. In addition, in order to solve the problem of slow training of DRL based on policy function in complex environments, the algorithm is parallelized, which provides the possibility for the practical application of DRL. In the latest research, PPO and A3C algorithm have more applications.

B. DRL-based Mapless Navigation

At present, the visual navigation without maps based on DRL is mainly navigated by the image from the first perspective of the robot and target information [19]. The information about the target is mostly the position or the image. Due to the small space and many markers in indoor navigation, the target image is often used as information of the target point [20]. The first perspective image observed by the target image and mobile robot currently observes as the input of the visual navigation model, the movement of the mobile robot is used as the output to realize end-to-end visual navigation [21]. In addition, in order to reduce the training time and enable the robot to reach the target faster, some auxiliary tasks will be added to the visual navigation model, such as depth estimation, reward prediction, pixel control and others [22]. In the existing research, there are fewer navigation research on complex environments. Due to the sparse reward problem in the DRL algorithm, it is difficult for mobile robots to gain experience [23]. Therefore, it is a challenging problem to realize the visual navigation of mobile robots in a complex environment.

III. Methodology

In this section, we will describe the whole structure of visual navigation model proposed in large-scale complex environment for mobile robots. We first use the function to define the visual navigation problem to reduce the complexity of the problem. In the second part, we introduce the local region model, which has better navigation performance in the new and old targets through network structure and reward function optimization. In the third part, we introduce the structure of the entire visual navigation model in large-scale complex environments for mobile robots.

A. Visual Navigation Problem Definition [1]

In order to implement navigation tasks in large-scale complex environments, the visual navigation model takes the RGB-D information and the target point as input and the continuous action of the mobile robot as the output. The trained model can make the mobile robot reach the target in a collision free path in a large-scale complex environment, and the new target that has not been trained is reached according to the model inference. Therefore, this problem can be defined as:

$$V_t = f(I_t, T_t, V_{t-1}) \tag{1}$$

Where I_t is the RGB-D image observed by the mobile robot, T_t is the current mobile robot and target position, V_t is the linear and angular velocity of the mobile robot at this moment, f is the mapping function.

B. Local Region Visual Navigation Model

The local visual navigation model of mobile robot uses the RGB-D image obtained from the first perspective of mobile robots and the destination as the input, and the continuous motion of mobile robot as output to realize the task of end-to-end visual navigation without maps. By training toward the target, the parameters of the local visual navigation model are updated, and the proposed models are used to deduce the new goals that have not been trained. The success rate of new target arrivals is used to determine whether the local visual navigation model continues to be updated. The higher the success rate of new target arrivals, the higher the degree of visual navigation model mastering environmental information.

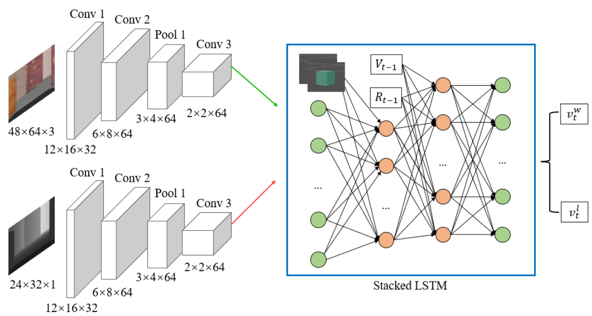


Fig. 1. local region visual navigation model.

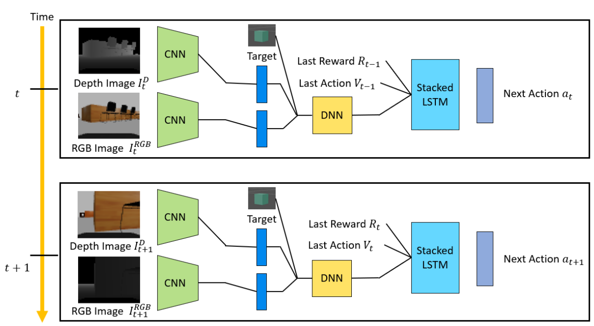


Fig. 2. local region model training.

a) network structure: The local region model is shown in Fig.1. The RGB image size is 48×64×3, and the depth image size is 24×32×1. The embedded vectors are obtained through the convolutional neural network (CNN); the speed of the mobile robot is V_{T-1}; the reward is R_{T-1}. The target position are integrated into the stacked LSTM. Then the action of the robot is determined according to the PPO algorithm. The green block represents the target information. The training process of the local region model as shown in Fig.2.

b) reward function: The design of the reward function is to make the mobile robot reach the target position quickly, so the design of the reward function is crucial. We design the reward function of collision, mobile robot speed, distance between the mobile robot and the destination, running time and other factors. When the mobile robot collides, it will be given a negative reward r_c; when the distance is less than c_d, it is determined that the mobile robot has reached the target position and get the reward r_a; in other cases, the mobile robot get rewards related to time and speed, d_t is the distance between the robot and the destination at time t, v_t^l, v_t^a is the linear velocity and angular velocity of the mobile robot, c_r, c_l ,c_a and c_t are model parameters. Therefore, the reward function of the mobile robot visual navigation model is defined as in (2):

$$f(x)=\begin{cases} r_c, & collision \\ r_a, & arrival \\ c_r\left(d_{t-1}-d_t\right)+c_l v_t^l+c_a\left(v_t^a\right)^2+c_t & otherwise \end{cases} \tag{2}$$

C. Distributed DRL based Visual Navigation [11]

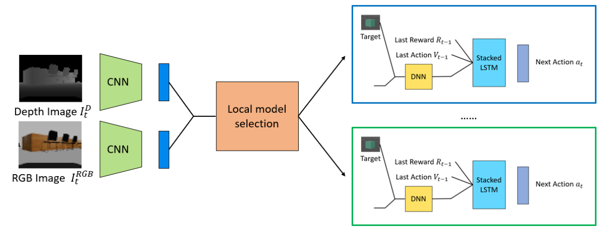


Fig. 3. distributed deep reinforcement learning based visual navigation.

Distributed deep reinforcement learning based visual navigation uses the training results of local models to complete the visual navigation in large complex scenes. First of all, the complex environment is divided according to the scene and then we train local models in each region. We use CNN to obtain the current state RGB-D image as the input and generate an embedded vector. According to the embedded vector and the target position, we determine the regions that need to be travelled and its order, and navigate in the local model until it reaches the target in the entire complex environment. The model we proposed is shown in Fig.3. Because it has fewer factors than the complex global model, it has high efficiency and speed.

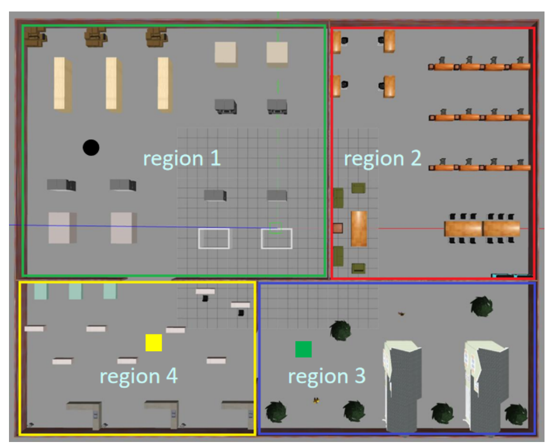
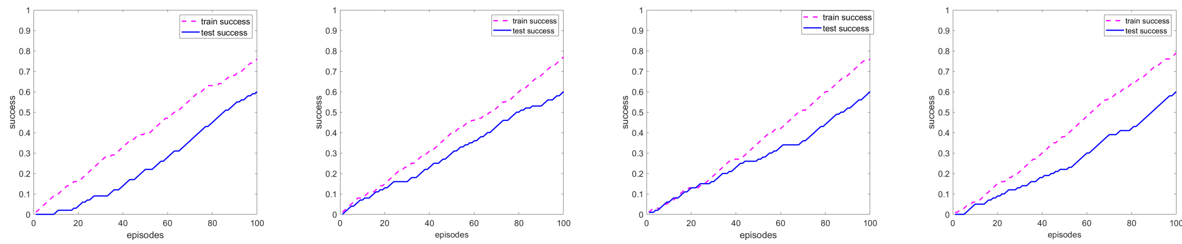


Fig. 4. large-scale complex environment.

IV. SIMULATION EXPERIMENTS

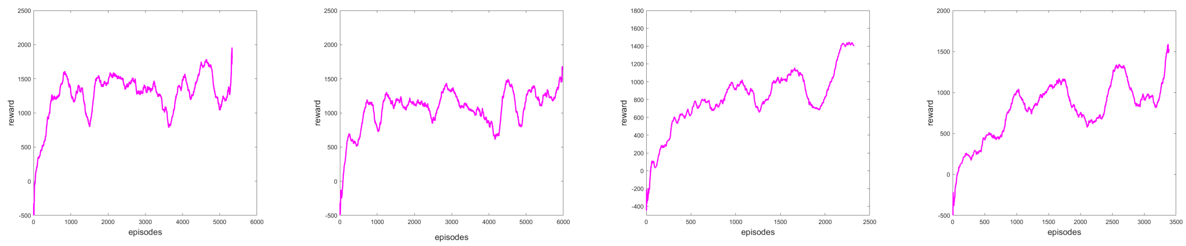
A. Simulation Environment

The simulation environment is established by Gazebo in Ubuntu16.04, as shown in Fig.4. The black circle represents the starting position of the mobile robot. The green cube is the



(a) local region 1 (b) local region 2 (c) local region 3 (d) local region 4

Fig. 5. local region model success rate.



(a) local region 1 (b) local region 2 (c) local region 3 (d) local region 4

Fig. 6. local region model reward over episodes.

training target to update the local model. The yellow cube is the reasoning target used to judge the model's understanding of the environment. The mobile robot is installed with a visual sensor to receive a real-time RGB-D image from the first perspective of the robot. The simulation factory environment consists of four regions: the production line, the office, the outdoor area, and the restaurant. There are channels between adjacent regions. For example, to reach region 3 from region 1, it is necessary to reach region 2 from region 1, and then from region 2 to region 3.

B. Result and Analysis

Compared with the success rate of navigation tasks of the last 100 episodes between the old and new targets in the local region model, as shown in Fig.5. When the success rate of the mobile robot reaching the new goal is 60%, the success rate of the old target reached more than 75%. As shown in Fig.6, the rewards that mobile robots obtained in the environment during the training process continued to rise.

The navigation success rate of visual navigation based on distributed DRL in local regions and the entire large-scale complex environment is shown in Table.I. The success rate of navigation of local region 1 is 95%, that of local region 2 is 77%, that of the local region 3 is 78%, that of local region 4 is 86%, and the navigation success rate in the entire complex scene is 84%. The navigation path from the starting point to the target is shown in Fig.7.

TABLE I
VISUAL NAVIGATION MODEL SUCCESS RATE

success rate				
region 1	***region 2***	***region 3***	***region 4***	***entire env***
95%	***77%***	***78%***	***86%***	***84%***

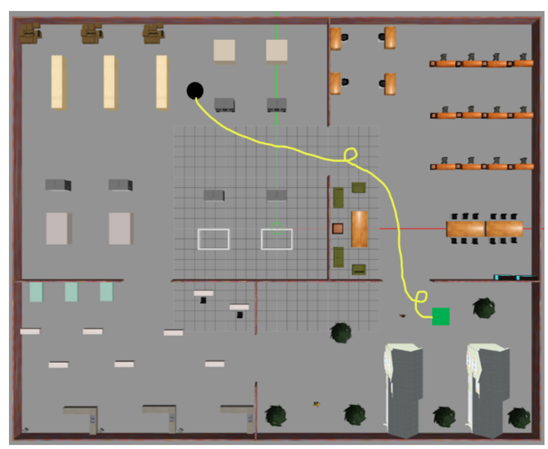


Fig. 7. visual navigation route.

V. CONCLUSION

In this paper, we improved the existing DRL visual navigation framework, and proposed a visual navigation model based on distributed DRL in large-scale complex environments. Firstly, the complex scene is divided into several regions. The local region models are trained by the RGB-D image and target position observed by the mobile robot. Then we select the local region model according to model selection until the target is reached. The success rate of the mobile robot reaching the target is 84% in the large-scale complex environment of GAZEBO, which has good navigation performance.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant number U20A20283 and U1813222, the Shenzhen International Cooperation Research Program under Grant number GJHZ20200731095009029, and the Department of Science and Technology of Guangdong Province under Grant number 2019TQ05Z654.

REFERENCES

[1] L. Ma, Y. Liu, J. Chen, and D. Jin, "Learning to navigate in indoor environments: from memorizing to reasoning," *arXiv preprint arXiv:1904.06933*, 2019.

[2] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "A brief survey of deep reinforcement learning," *arXiv preprint arXiv:1708.05866*, 2017.

[3] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part i," *IEEE robotics & automation magazine*, vol. 13, no. 2, pp. 99–110, 2006.

[4] S. S. Ge, X. Lai, and A. A. Mamun, "Boundary following and globally convergent path planning using instant goals," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 35, no. 2, pp. 240–254, 2005.

[5] W. Rone and P. Ben-Tzvi, "Mapping, localization and motion planning in mobile multi-robotic systems," *Robotica*, vol. 31, no. 1, pp. 1–23, 2013.

[6] E. Nwaonumah and B. Samanta, "Deep reinforcement learning for visual navigation of wheeled mobile robots," in *2020 SoutheastCon*. IEEE, 2020, pp. 1–8.

[7] L. C. Garaffa, M. Basso, A. A. Konzen, and E. P. de Freitas, "Reinforcement learning for mobile robotics exploration: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[8] F. Zeng, C. Wang, and S. S. Ge, "A survey on visual navigation for artificial agents with deep reinforcement learning," *IEEE Access*, vol. 8, pp. 135 426–135 442, 2020.

[9] Q. Fang, X. Xu, X. Wang, and Y. Zeng, "Target-driven visual navigation in indoor scenes using reinforcement learning and imitation learning," *CAAI Transactions on Intelligence Technology*, 2021.

[10] K. Zhu and T. Zhang, "Deep reinforcement learning based mobile robot navigation: A review," *Tsinghua Science and Technology*, vol. 26, no. 5, pp. 674–691, 2021.

[11] S.-H. Hsu, S.-H. Chan, P.-T. Wu, K. Xiao, and L.-C. Fu, "Distributed deep reinforcement learning based indoor visual navigation," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 2532–2537.

[12] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," *arXiv preprint arXiv:1312.5602*, 2013.

[13] H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double q-learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 30, no. 1, 2016.

[14] Z. Wang, T. Schaul, M. Hessel, H. Hasselt, M. Lanctot, and N. Freitas, "Dueling network architectures for deep reinforcement learning," in *International conference on machine learning*. PMLR, 2016, pp. 1995–2003.

[15] V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu, "Asynchronous methods for deep reinforcement learning," in *International conference on machine learning*. PMLR, 2016, pp. 1928–1937.

[16] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[17] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *International conference on machine learning*. PMLR, 2015, pp. 1889–1897.

[18] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[19] Y. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, and A. Farhadi, "Target-driven visual navigation in indoor scenes using deep reinforcement learning," in *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2017, pp. 3357–3364.

[20] X. Wang, Q. Fang, and X. Xu, "Target-driven indoor visual navigation using inverse reinforcement learning," in *2020 International Conference on Image, Video Processing and Artificial Intelligence*, vol. 11584. International Society for Optics and Photonics, 2020, p. 115841R.

[21] L. Kästner, C. Marx, and J. Lambrecht, "Deep-reinforcement-learning-based semantic navigation of mobile robots in dynamic environments," in *2020 IEEE 16th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2020, pp. 1110–1115.

[22] J. Kulhánek, E. Derner, and R. Babuška, "Visual navigation in real-world indoor environments using end-to-end deep reinforcement learning," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4345–4352, 2021.

[23] S. Nilwong and G. Capi, "Outdoor robot navigation system using game-based dqn and augmented reality," in *2020 17th International Conference on Ubiquitous Robots (UR)*. IEEE, 2020, pp. 74–80.