Target Search Based on Scene Priors

Shengyang Lu, Lanjun Liang, Huailin Zhao✉, Fangbo Zhou, Feng yao
School of Electrical and Electronic Engineering,
Shanghai Institute of Technology, Shanghai, China
E-mail: 405838753@qq.com, zhao_huailin@yahoo.com
www.sit.edu.cn

Abstract

Aiming at the problems of reinforcement learning algorithm in target search tasks, such as low accuracy and low fault tolerance, this article mainly introduces a method of reinforcement learning target search based on scene prior in simulation environment. This method mainly uses graph convolutional neural network to extract the current object relationship as the input of prior knowledge. Secondly, it uses the actor-critic algorithm to take the agent's vision, position and prior knowledge as input to decide the agent's next navigation. Finally, use path planning to navigate to the target point to find the target. Through experiments conducted in Habitat and compared with the previous algorithm, the experiment shows that this method is better than the previous algorithm in target search accuracy and navigation efficiency.

Keywords: Target search, Reinforcement Learning, Scene priors, visual navigation

1. Introduction

In recent years, the field of robot research has been committed to expanding the ability of robots to explore the environment, understand the environment, interact with the environment and communicate with people. As one of the most important tasks of agent visual navigation, visual semantic navigation requires the robot to navigate to the target location by using the observed visual information according to the given target in the unknown environment. The agent solves two problems at the same time. The first is where to go, that is, where the target object is, and the second is how to get there, that is, planning an effective path to navigate to the target object. This has far-reaching significance for both the research in the field of artificial intelligence and the application in real life, such as disaster and battlefield rescue, smart home, unmanned driving and so on. Moreover, it is also of far-reaching significance for the research in other fields, such as embodied question answering [1], visual semantic navigation [2], and Visual Dialogue navigation [3].

The traditional navigation method [4] usually uses the environment map for navigation, and divides the navigation task into three steps: drawing map, positioning and path planning. This method usually needs to build a 3D map in advance, as well as reliable map positioning and path tracking. However, the map can not be used in the invisible environment. Recently, the success of data-driven machine learning strategies for various control and perception problems has opened up a new way to overcome the limitations of the previous method [4]. The key point of these methods is to directly learn the mapping between the original observation and operation of the end-to-end way of the task. These methods take advantage of the ability of previous navigation experience in a new similar environment, whether there is a map or not. Reinforcement learning (RL) is often used in visual navigation. However, reinforcement learning still has some problems, such as low generalization ability, low navigation efficiency and low accuracy.

Recently, chaplot [5] proposed a target driven navigation method based on learning model and won the first place in CVPR 2020 habitat objectnav challenge, called "goal oriented semantic exploration" "Visual navigation using semantic mapping and reinforcement learning can cover the whole scene for target search to the greatest extent. However, this method still has the problems of low learning efficiency and low navigation accuracy.

However, human navigation still exhibits excellent generalization performance, which cannot be explained by spatial or topological memory. For example, people who come home for the first time will naturally go to the kitchen (instead of outdoor or toilet) to pick up plates; From the kitchen to the bedroom, they know that the living room may be a midway point. Although visually different, this kind of semantic knowledge, the "close" relationship of semantic entities, is naturally shared across environments and can be learned from past experience to guide future navigation.

Prior knowledge can not only help to navigate to known objects, but also help agents find targets based on the current visual exploration of unknown environment. Based on this, we propose a target driven navigation model based on semantic prior. The model uses semantic prior knowledge to assist navigation, and improves the ability of agent target search by learning the relationship between objects power.

In order to evaluate our model, we trained and tested our model in the habitat [6] simulation environment. Habitat simulation environment allows us to train AI agents in realistic and efficient 3D simulators. The simulation environment includes Gibson [7] and matterport3d [8] data sets, including "kitchen", "toilet", "living room" and other scenes, allowing agents to explore the whole room to find targets, which greatly increases the difficulty of the task. The agent is initialized to the random position in the scene, and explores through visual information to find the target. Experiments show that our target search algorithm based on scene a priori is superior to other methods in both navigation success rate and average step.

In this paper, our innovations can be summarized as follows:

- We propose a target search algorithm based on semantic a priori. The algorithm uses semantic a priori knowledge to assist the agent to explore the unknown environment, and helps the agent to search the target through the relationship between objects.
- We train and test our method in the habitat simulation environment. Experiments show that our method is superior to other algorithms in both the success rate of target search and the efficiency of navigation.

2. Related Work

Navigation is one of the most fundamental issues for mobile robots. Traditional methods such as slam construct metric maps by perceiving signals, which are subsequently used in planning. Recently, due to advances in deep learning, end-to-end methods have been applied to navigation in several fields, such as mazes, indoor scenes [24,25], autonomous driving [26,27]. There is also an excellent summary of recent progress [2,37]. We focus on indoor navigation scenarios, using the house3d environment [28], which contains relationships between semantic entities that are consistent with the real world and provides ground truth labels of objects and scenarios.

2.1. Semantic navigation

Semantic navigation and target search tasks have far-reaching significance for both the development of artificial intelligence and the application of daily life. Recently, many methods use reinforcement learning as a navigation decision module to control agents to navigate in unknown environments. Such as A3C [9], PPO [10], DQN [11], etc. In addition, Liang [12] et al. Proposed an algorithm to explicitly model the scene a priori using the confidence perception semantic scene completion module to complete the scene and guide the navigation planning of the agent. Moreover, many opportunistic memory mechanisms[13],attention mechanisms [14] and algorithms of transfer learning [15] and imitation learning [16] are also applied to visual navigation. Although a large number of algorithms have improved the performance of searching objects in agent exploration environment, they still face the characteristics of small target search range, low accuracy and low robustness. And these algorithms do not take into account the relationship characteristics between objects.

2.2. Scene priors

The target search algorithm based on scene a priori takes more account of the relationship between objects, and uses this relationship to assist agents in environment exploration and target search. In recent years, many people use prior knowledge to assist various tasks, and have made good achievements in such tasks. For example, scene atlas is used for traffic light search [17], road detection [18] and dialogue tasks [19]. Moreover, in terms of visual navigation, Yang [20] et al. Proposed a reinforcement learning model framework that uses graph convolution neural network coding to assist exploration of scene prior knowledge. Liu [21] and others applied the scene prior knowledge to the multi-agent system, which effectively improved the efficiency of multi-agent system navigation and target search. Li [22] et al. Combined prior knowledge with meta learning and made great progress in map free visual navigation.

3. Problem description

In unknown scene s, The visual information of the agent is recorded as O, $O=\{O_t^1,O_t^2,...,O_t^k\}$. At the beginning, the agent is initialized to the random position and random posture in the room, which is recorded as $P_0=\{x_0,z_0,\theta_0\}$. (x_0,z_0) Represents the random position of the agent initialized in the room at the initial time, θ_0 is the initial rotation angle. Initially, the agent receives the tag T of the target as input . At each time point, the agent receives visual information from the visual sensor O_t, and its odometer receives attitude information P_t. Visual information O_t consists of first person RGB and depth information.

At each time point t, agent A receives the first person visual information O_t from the visual sensor, obtains the pose information P_t from the odometer. Visual information and pose information are used to predict the corresponding semantic map m_t. Each agent learns its corresponding navigation strategy π to determine the corresponding expected navigation point L, $L_t=\pi(m_{t-1},T,L_{t-1})$. In order to reach the expected target point, the agent obtains the next action A_t through path planning. When action A_t is performed in each agent, its visual input is updated to O_{t+1} and the odometer input P_{t+1}. Go back and forth until you find all the targets.

4. Problem description

This paper proposes an actor critic target search model based on scene a priori, which is mainly composed of four parts. As shown in Figure 1, they are semantic mapping module, prior knowledge extraction module, feature fusion module and action decision module. The semantic decision module maps the first person vision into top-down semantic map features. The object relationship feature uses the scene prior knowledge extraction module to extract the object relationship in the first person RGB information. The feature fusion module is to fuse the current object feature relationship and semantic mapping vector with the previous state. The action decision module is to decide the next expected goal and generate the next action that the agent needs to perform.

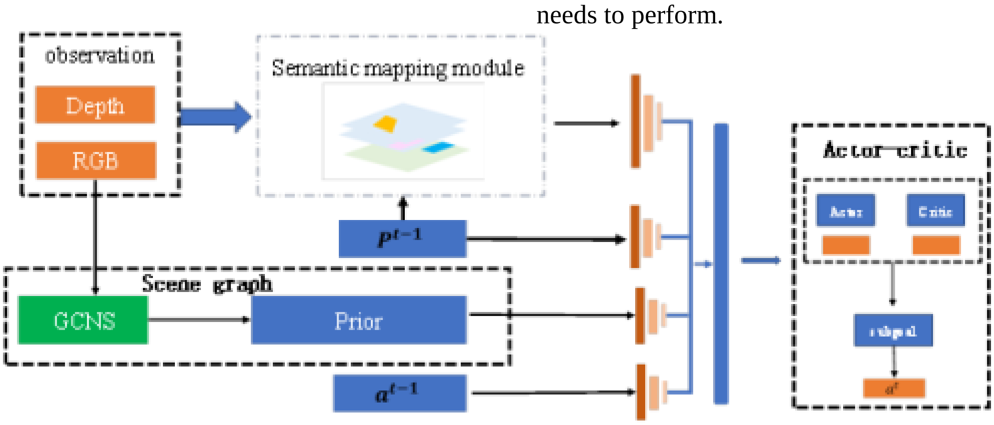


Fig. 1: overall framework

4.2.Semantic mapping module

In this module, the method similar to literature [5] is mainly used to generate idiom semantic mapping map. Our agent uses RGB and depth to generate point cloud information, voxels it, and finally turns it into a semantic map. Then, reinforcement learning is used to generate the expected goals of each agent. In this process, semantic maps are constantly being improved.

We mainly use spatial semantic maps and can access the location of agents at any time. The structure of spatial semantic map is a matrix of $K \times M \times N$, in which the size of $M \times N$ is the size of the actual map. Firstly, the size of $K \times 240 \times 240$ is generated, then the matrix is mapped to the size of the actual map, and finally the spatial semantic map m_t is generated. The semantic map is composed of voxels, and the size of each voxel block is $16m^2$. The number of channels of spatial semantic map k = 2 + C, C represents the semantic information of each object in the actual environment, and the first two layers represent the map size and explored places respectively. Each element on the map represents the place that has been explored and corresponds to the position of the semantic channel object in the map. During initialization, the map is initialized with an all 0 matrix, such as $[0]^{k \times M \times N}$, and the agent is initialized in a random position

4.3. Prior knowledge extraction

The object relationship feature extraction module adopts a method similar to that in literature [31], brings semantic knowledge into the reinforcement learning framework, integrates the relationship features between current environmental objects using graph convolution networks (GCNs) [32], and dynamically updates and saves them when the agent receives the environmental information.

With the constant navigation of the agent, the agent's understanding of the whole environment is increasingly clear, and the relationship between the object and the object is gradually improved. These relational vector assistant agents are used to find the target objects. Based on these relational agents, they can find objects more efficiently and with a higher success rate.

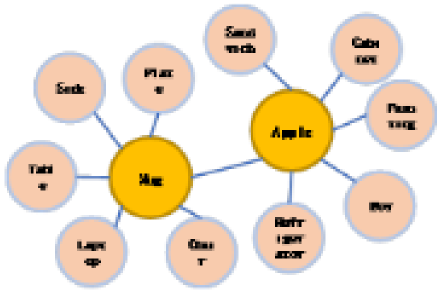


Fig 2. Scene priors

In this paper, the prior knowledge of the scene is represented in the form of undirected graph $G=\{V,E\}$.The node in V represents different types of objects, and the edge E represents the special positional relationship between the two types of objects.

As an extension of graph neural network to graph structure, GCNs aims to learn the functional representation of a given graph $G=\{V,E\}$. We generalize all nodes into characteristic matrices. The graph structure is represented by binary adjacency matrix $F=[F_1, F_2 ... F_{|V|}]$. we standardize the matrix A to obtain the matrix $\hat{A}$. Each node of GCNs output is represented as $Z=[z_1, z_2, ..., z_{|V|}]$. So we can get:

$$H^{(l+1)} = f(\hat{A}H^{(l)}W^{(l)})$$

Where, $H^{(0)} = X$, $H^{(L)} = Z$, $W^{(l)}$ is the parameter of layer l and L is the total number of layers of GCNs.

In this section, three layers of GCNs are used, the input is RGB image, the output of the first two layers is 1024 dimensional potential features, and the output of the last layer is the value output by each node to obtain the feature vector $|V|$.

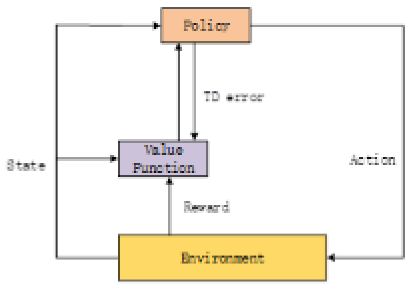


Fig 3. Navigation model

The feature vector is the semantic coding information of the current scene and environment context. Finally, this eigenvector is mapped to a 512 dimensional eigenvector $f_{K,t}{}^{k}$. The 512 dimensional feature vector is used as the input feature to assist the agent in visual navigation.

4.3. Navigation decisions

At this stage, we mainly use reinforcement learning to solve the problem of our multi-agent visual navigation. The reinforcement learning model is shown in the figure below. Our long-term goal is to let the agent team find the goal of concern, so the agent must navigate to the area where the goal is located. The multi-agent spatial semantic map described in Section 4.2 is used as the input, and the short-term goals that can be reached by each agent are generated through the network above. The output short-term goals are in the form of {x, z}. The short-term goals will lead the agent team to conduct visual guidance in an interactive environment. When the agent sees the target concerned by the problem, the channel corresponding to the target in the spatial semantic map will be non-0. When the channel corresponding to the target in the spatial semantic map is 0, the short-term goal will guide the agent to explore more places until the information of the spatial semantic map is enough to answer the problem. As shown in Fig 3.

5. Experimental setup

We use habitat simulation environment to test our model. Haibitat contains Gibson and matterport3d data sets, but we only choose Gibson data set as simulation environment to train and test the effect of our model.

The agent team is initialized at any position in the environment, so the observation space is composed of the first person vision (RGB and depth) of each agent, which is a 4 * 480 * 640 matrix. The coordinates of the agent mainly include the position and attitude information of the agent, which is composed of $\{x, z, \theta\}$, which is a 1×3 matrix. The action space of agent is continuous. When the agent finds the target, that is, there is a semantic mapping of the target label in the semantic map, it means that the navigation is successful, and then the navigation will end, or the length of the action sequence of the agent exceeds 3000, then we think that the agent is very familiar with the scene this time, can not find the target, and there is no other significance to continue searching, Therefore, it will also end this navigation and determine that the navigation is wrong.

In order to take more account of the relationship between objects, the semantic map we designed includes 15 categories. During the retraining process of our hierarchical reinforcement learning network, the reward is saved every 200 actions, the batch size is 36 in each reinforcement learning update weight process, and it is updated every 4 epochs. We use the Adam optimizer with a learning rate of 0.000025, and the weight size is set to γ=0.95.

6. Experimental setup

This article uses two evaluation functions: success rate (SR) and average path length (SPL) to test the effect of the algorithm. The success rate is defined as:

$$SR=\frac{1}{N_{task}}\sum_{i=1}^{N_{task}} R_i \qquad (2)$$

Among them, when the i round of experiment is successful, $R_i = 1$ otherwise $R_i = 0$, N_{task} is the total number of rounds of the experiment. The higher the success rate, the better the search effect of agents.

The average path length is defined as:

$$SPL=\frac{1}{N}\sum_{i=1}^{N} S_i \frac{l_i}{\max(p_i, l_i)} \qquad (3)$$

Among them p_i is the length of the path that the agent moves in the first round, l_i is the shortest path length from the initial position of the agent to the target, and N represents the total number of rounds. S_i defines whether this round is successful. The lower the average path length, the higher the efficiency of multi-agent searching for targets.

Distance to succes is defined as:

$$DST=\max(\|\mathrm{x}_T - G\|_2 - d_s, 0) \qquad (4)$$

Where $\|\mathrm{x}_T - G\|_2$ is the L2 distance of the agent from the goal location at the end of the episode, d_s is the success threshold.

6.1. Navigation decisions

We use two end-to-end Reinforcement Learning (RL) methods as baselines:

RGBD + RL: A common recursive RL strategy initialized with the resnet18 trunk,. Proxy poses and target object categories are passed through the embedded layer and attached to the loop layer input.

RGBD + Semantics + RL: This baseline passes semantic segmentation and object detection prediction and rgbd input to recursive RL strategy. We use the pre trained mask RCNN, which is the same as the RCNN used in the proposed model, for semantic segmentation and object detection in this baseline. Rgbd observations are encoded using resnet18 backbone visual encoder, and proxy pose and target object are encoded using the above embedded layer.

Both RL-based baselines are trained using proximity strategy optimization [10], using a dense reward that reduces the distance to the nearest target. We designed two additional baselines based on a combination of goal-independent exploration methods and heuristic-based local goal-driven strategies.

SemExp[5] : This baseline maps the visual inputs onto a three-dimensional semantic map, predicting the semantic map by using a method similar to the active nerual slam. Meanwhile, semantic map information is coded as vectors and trained using PPO's algorithm, and a navigation strategy is learned.

Our method mainly improves on this algorithm, adds scene prior knowledge on the basis of semantic map, and codes as knowledge vector to fuse with the semantic map, puts it into the network for learning, and finally learns a single-agent navigation strategy.

6.2. Quantitative experiment

TABLE1 Comparison of quantitative experimental results

	SPL	SR	DTS
Random	0.004	0.004	3.893
RGBD + RL	0.027	0.082	3.310
RGBD + Semantics + RL	0.049	0.159	3.203
SemExp	0.199	0.544	1.723
ours	**0.304**	**0.621**	**0.555**

We trained our algorithm over 10,000 times on Gibson and evaluated its performance in our test set. As shown in the table above, our algorithm has the highest performance compared to several algorithms, whether SPL, SR or DTS. By analyzing the table above, we can draw a few conclusions：

For random walk, the scene is larger in Gibson simulation environment, and the agent needs to navigate to find the target. For these three criteria, random walk is the worst. For RGBD+RL, RL models the process of exploration as a partially visible Markov model, but the model can not fit a navigation model perfectly due to less information, and the training speed is slow.

RGBD + semantic + RL and semexp take more into account the semantic information and spatial information between scenes, so they have a certain improvement in accuracy. However, they do not take into account the relationship characteristics between objects, so our model is superior to other methods in the above three indicators. Therefore, we draw a conclusion, Our model is more suitable for semantic exploration in unknown environment.

6.2. Qualitative experiment

In this section, we visualize first person vision for intelligent navigation as well as semantically predicted maps and navigation paths, with the specific structures shown in figure 4.

In this experiment, given an intelligent one semantically tagged as a " toilet " intelligent achieves this goal by predicting a semantic map and navigating, eventually finding a target, when it is found, turning the color of the target blue, providing us with the maximum priority that we need to navigate, and then navigating directly to the location of that target.

As shown in figure4, at the beginning the intelligent is randomly initialized at random locations throughout the environment, and it is not clear to the intelligent what the environment looks like, so the whole semantic map is blank. As the intelligence receives first person visual information, the intelligence begins to predict the semantic map, predict the situation across the scene, and compute an intended target based on the current incomplete semantic map, as well as scene prior information, which is where the intelligence wants to go. As shown in figure 4, at the beginning the intelligent is randomly initialized at random locations throughout the environment, and it is not clear to the intelligent what the environment looks like, so the whole semantic map is blank.

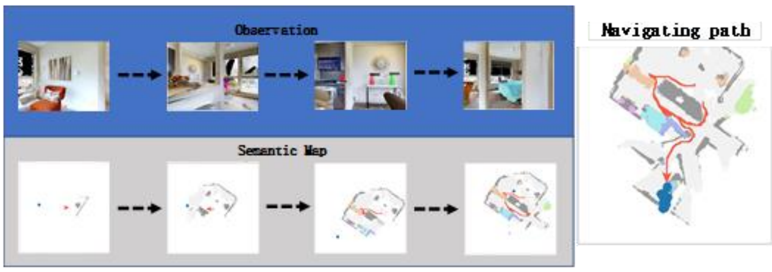


Fig 4.visual semantic navigation path

As the intelligence receives first person visual information, the intelligence begins to predict the semantic map, predict the situation across the scene, and compute an intended target based on the current incomplete semantic map, as well as scene prior information, which is where the intelligence wants to go. In the algorithm design, this design takes LK optical flow tracking as an extension of the better key point. Therefore, the LK optical flow tracking node can be used to subscribe the key points that have been obtained in the better key point node. The algorithm can be used to pass the current key point. Two grayscale pictures predict the next set of key points and perform backward prediction.

5.Conclusion

This paper mainly introduces a visual navigation method based on scene a priori. Firstly, this method uses GCNs to calculate the relationship between objects and encode them. Secondly, the method calculates the prediction semantic map of the scene through the first person visual information and its own pose information. Encode the above information, fuse features, and conduct visual navigation through actor critical algorithm to finally find the target. Through qualitative and quantitative experiments, it is concluded that the algorithm proposed in this paper is effective in both success rate and average path length.

6.References

1 Das, Abhishek, et al. "Embodied question answering. " *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.

2 Anderson P, Wu Q, Teney D, et al. Vision-and-language navigation: Interpreting visually-grounded navigation instructions in real environments[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 3674-3683.

3 Thomason J, Murray M, Cakmak M, et al. Vision-and-dialog navigation[C]//Conference on Robot Learning. PMLR, 2020: 394-406.

4 Bailey T, Durrant-Whyte H. Simultaneous localization and mapping (SLAM): Part II[J]. IEEE robotics & automation magazine, 2006, 13(3): 108-117.

5 Chaplot D S, Gandhi D P, Gupta A, et al. Object goal navigation using goal-oriented semantic exploration[J]. Advances in Neural Information Processing Systems, 2020, 33.

6 Savva M, Kadian A, Maksymets O, et al. Habitat: A platform for embodied ai research[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019: 9339-9347.

7 Xia F, Zamir A R, He Z, et al. Gibson env: Real-world perception for embodied agents[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 9068-9079.

8 Chang A, Dai A, Funkhouser T, et al. Matterport3d: Learning from rgb-d data in indoor environments[J]. arXiv preprint arXiv:1709.06158, 2017.

9 Babaeizadeh M, Frosio I, Tyree S, et al. GA3C: GPU-based A3C for deep reinforcement learning[J]. CoRR abs/1611.06256, 2016.

10 Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[J]. arXiv preprint arXiv:1707.06347, 2017.

11 Hernandez-Garcia J F, Sutton R S. Understanding multi-step deep reinforcement learning: a systematic study of the DQN target[J]. arXiv preprint arXiv:1901.07510, 2019.

12 Liang Y, Chen B, Song S. Sscnav: Confidence-aware semantic scene completion for visual semantic navigation[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 13194-13200.

13 de Andrade Santos I B, Romero R A F. Deep Reinforcement Learning for Visual Semantic Navigation with Memory[C]//2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE). IEEE, 2020: 1-6.

14 Seymour Z, Thopalli K, Mithun N, et al. MaAST: Map Attention with Semantic Transformersfor Efficient Visual Navigation[J]. arXiv preprint arXiv:2103.11374, 2021.

15 Wang L, Zhao L, Huo G, et al. Visual semantic navigation based on deep learning for indoor mobile robots[J]. Complexity, 2018, 2018.

16 Wang X, Huang Q, Celikyilmaz A, et al. Reinforced cross-modal matching and self-supervised imitation learning for vision-language navigation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 6629-6638.

17 Barnes D, Maddern W, Posner I. Exploiting 3D semantic scene priors for online traffic light interpretation[C]//2015 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2015: 573-578.

18 Alvarez J M, Gevers T, Lopez A M. 3D scene priors for road detection[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2010: 57-64.

18 Mishra A, Alahari K, Jawahar C V. Scene text recognition using higher order language priors[C]//BMVC-British Machine Vision Conference. BMVA, 2012.

19 Yang W, Wang X, Farhadi A, et al. Visual semantic navigation using scene priors[J]. arXiv preprint arXiv:1810.06543, 2018.

20 Liu X, Guo D, Liu H, et al. Multi-Agent Embodied Visual Semantic Navigation with Scene Prior Knowledge[J]. arXiv preprint arXiv:2109.09531, 2021.
21 Li F, Guo C, Luo B, et al. Multi goals and multi scenes visual mapless navigation in indoor using meta-learning and scene priors[J]. Neurocomputing, 2021, 449: 368-377.
22 He K, Gkioxari G, Dollár P, et al. Mask r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.
23 Wang L, Zhao L, Huo G, et al. Visual semantic navigation based on deep learning for indoor mobile robots[J]. Complexity, 2018, 2018.
24 Liang Y, Chen B, Song S. Sscnav: Confidence-aware semantic scene completion for visual semantic navigation[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 13194-13200.
25 Wang J, Liu J, Kato N. Networking and communications in autonomous driving: A survey[J]. IEEE Communications Surveys & Tutorials, 2018, 21(2): 1243-1274.
26 Yurtsever E, Lambert J, Carballo A, et al. A survey of autonomous driving: Common practices and emerging technologies[J]. IEEE access, 2020, 8: 58443-58469.
27 Panayi N. In-house 3D designing of lingual and labial customized brackets using composite-made customized bases[J].

Authors Introduction

Mr. Shengyang Lu

He received the B.S. degree from Shanghai Institute of technology, Shanghai, China, in 2018. He is currently pursuing the M.S. degree in Control Science and Engineering at Shanghai Institute of Technology. His main research interests are deep learning and ROS robots.

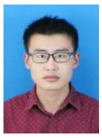
Ms.Lanjun Liang

She received B.S. degree in automation from Shanghai Polytechnic University, Shanghai, China, in 2020. She is studying for the master degree of control science and engineering at Shanghai Institute of Technology, Shanghai, China. her main research interests are deep learning and ROS robots.

Dr. Huailin Zhao

He received his PhD from Oita University, Japan in 2008. He is a professor in the School of Electrical & Electronic Engineering, Shanghai Institute of Technology, China. His main research interests are robotics, multi-agent system and artificial intelligence. He is the members of both IEEE and Sigma Xi.

Mr. Fangbo Zhou



He received the B.S. degree in automation from Chuzhou University, Chuzhou, China, in 2019. He is currently pursuing the M.S. degree in safety engineering at Shanghai Institute of Technology, Shanghai, China. His research interests include crowd counting and ROS robots.

Mr.Feng Yao

He received the B.S. degree from Shanghai Institute of technology， Shanghai, China, in 2020. He is currently pursuing the M.S. degree in Control Science and Engineering at Shanghai Institute of Technology, Shanghai, China. His main research interests are reinforcement learning and human-robot collaboration.