Method of Indoor Navigation and Safety Improvement by SE-Dense-DQN

Jie Zhou¹, Xin Zheng² and Mengting Chen³,*

¹,²,³Faculty of Information and Technology, Beijing University of Technology, Beijing, China

*Corresponding author email: cmt2001@emails.bjut.edu.cn

Abstract. Since the increasing requirement of indoor navigation with the assurance of safety, Deep Learning was been used in this area by numerous scholars. In order to help unmanned aerial vehicle navigate in various environment without prior experience, this paper proposed the method of Deep Q Learning with the Squeeze-and-Excitation Networks and Dense blocks. The structure of the Q-Network also amended with the information of LIDAR processed by Long Short-Term Memory to make sure the accuracy. The structure of the network shows in figure 6. The experiment conducted in simulated environment. Experimental results indicated that this method excelled in efficiency and accuracy than other classic Deep Q Networks like classic Deep Q Network and Double Deep Q Network. The results displaying in figure 8 and figure 9 shows unmanned aerial vehicle can arrive at target place more accurately than other networks.

Keywords: Indoor navigation; DQN; LSTM.

1. Introduction

With the development of UAV (Unmanned Aerial Vehicle), the requirement of navigation indoor without previous experience is becoming increasing. In order to enhance the adaptability of indoor UAVs, researches have turned to the field of improving the real-time and accuracy of UAVs. Deep Reinforcement Learning is one of the solutions to access to the aim. DQN (Deep-Q Learning), which learns from the progress of exploration can meet the requirements of indoor navigation.

The biggest challenge of the navigation and safety in unfamiliar environments is the lack of the known information. For computer which controls the UAV, the more information computer gets, the more accurate which state of the agent it can analyse. However, collecting datasets in the environment takes lots of time, which makes algorithm like Simultaneous Localization and Mapping (SLAM) cannot fit various environments even though it is accurate and safe with the optimize by scholars [1][2]. The next state of the system is determined by the present state and the action which adopted by agents according to the Markov Decision Process (MDP) which is the basic of the Reinforcement Learning. As a result, in terms of DQN, the decision can be made just with the present information and the valuation of the possible actions that might be taken. The preparation of samples of indoor scenes is unnecessary, for the states of agent are accumulates during the exploration which can be used in the training of Value Function. These characteristics hardly satisfies the navigation in the unfamiliar environments.

In recent researches, improving real-time and accuracy of indoor navigation is the hot topic. Methods of the improvement can be divided into two categories: pre-processing and network. For pre-processing, the methods can be summarized as reduction of information. Normalization of images
[3][4][5], Cropping or shrinking images [5][6] and the usage of Information Theory [7] are all the methods. Even though the normalization and cropping can make the image processing faster, lot of information lost by these two operations. The usage of information theory can reduce the work of network in handling the data while the work of pre-processing is larger than other methods.

In the field of path planning and obstacle avoidance using, most of the current network models are general models that can be used for outdoor environment (especially in urban environment) [8], and that there are many sensors needed, therefore the problem of sensor power exists. Option-based hierarchical reinforcement learning [9] and MIMO [10] can be used to solve this problem, but for indoor environment, it’s not efficient.

In An approach for UAV indoor obstacle avoidance based on AI technique with ensemble of ResNet8 and Res-DQN by Hai, the author proposed a method of indoor obstacle avoidance: combine the deep reinforcement learning and Res-Net. The main idea of Res-Net was to add more layers without degrading network performance, which resulted in low-cost performance for adding more layers, and a large number of parameters were underutilized.

In order to overcome the drawbacks above, this paper chooses the method of shrinking images by Bilinear Interpolation as the pre-processing method to reserve more image details and get shift-Invariance, whose calculation volume is much fewer. Dense-Net which is the optimization of Res-Net replaced the Res-Net module in Res-DQN and SE module was inserted in the Q Network. The Network optimized calls Squeeze Excitation-Dense-Deep Q Network (SE-Dense-DQN). Long Short-Term Memory (LSTM) was also used to merge the information of LIDAR with the information of images to improve the accuracy. The results of experiments verify the effect of the optimization. The main contribution of this paper is as follows:

(1) Improved the shift-invariance by zoom in image with the method of bilinear interpolation instead of pooling or convolution in Q Network.

(2) Compared with Res-DQN, this paper replaced the Res-Net with Dense-Net to mitigate the gradient disappearance, enhance the transmission of feature maps, make more effective use of features and reduce the number of parameters.

(3) SE-Net Module and LSTM were used in this paper to get more information of the agent’s state and the suppress features that are of use to the tasks of navigation indoor.

The main implementation block diagram is shown in figure 1:
usually pooling or stepwise convolution that are all sensitive to image panning in the classic CNN. In the research of Zhang’s, he pointed out that two box filters which equal to bilinear Interpolation can be applied in the down-sampling. In his paper, he did experiments on VGG16 and ResNet18. The results showed that consistency is higher that classic methods.

Current control methods using monocular images for UAV obstacle avoidance are heavily dependent on environment information. These controllers do not fully retain and utilize the extensively available information about the ambient environment for decision making. Singla’s method [12] is to add temporal attention to deep recurrent Q-network and obtain depth maps from RGB images through Conditional Generative Adversarial Network (CGAN). However, this paper didn’t optimize the pre-processing which may make the performance of the network some better. This point was also mentioned in the last section of the paper.

3. Methodology

3.1. Pre-processing by Bilinear Interpolation

Since DQN has the same process of convolution, the method of Bilinear Interpolation can be applied as well to get higher consistency and shift-invariance. The principle of bilinear interpolation is as follows:

Let the width of origin image as m, the height of it as n, Let the width of target image as a, and the height of it as b. When take any point of the target image \( P(x, y) \), the relative point in origin image is \( P' \left( \frac{m}{a}x \cdot \frac{n}{b}y \right) \). The coordinates of \( P' \) is usually radix. \( Q_{11}(x_1, y_1), Q_{12}(x_1, y_2), Q_{21}(x_2, y_2), Q_{22}(x_2, y_2) \) are all the points that near to the point \( P \). The value of the point that should be interpolate in the \( P \)‘s place needs to be calculated by the Formula below and the figure 2 shows the distribution diagram of each point:

\[
f(x, y) = \frac{f(Q_{11})}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y_2 - y) + \frac{f(Q_{21})}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y_2 - y) \\
+ \frac{f(Q_{12})}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y - y_1) + \frac{f(Q_{22})}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y - y_1).
\]

Figure 2. The Schematic of Bilinear Interpolation.

3.2. Indoor navigation by Deep Q Network

Reinforcement Learning (RL), which is a big category of machine leaning, can learn from what did and take actions according to what the state of the agents and how to maximize the value of reward function [13]. RL is comprised of four elements: a policy, a reward signal, an environment, and a utility function, and it’s a good candidate for resolving the high-complexity situations and capturing the realistic scenarios. Q-learning learns a strategy and tells the agent take the best action in the specific state. It doesn't need a model of the environment and can deal with the problems of random transition and reward without adaptation. Deep Q Network (DQN), proposed by DeepMind in 2013[14], is the combination of deep neural network and Q-learning. We can input the state and action,
and then get the Q value of the action taken with the neural network. In this way, we can directly use the neural network to generate the Q value.

DQN contains two neural networks with the same structure but different parameters: the network estimated by eval net used to predict $Q_{eval}$ has the latest parameters, but target net predicts $Q_{next}$ using parameters in the memory bank

$$Y^DQN_t = R_{t+1} + \gamma \max_a Q(S_{t+1}, a; \theta^T)$$

Target net is a historical version of eval net with a set of parameters long before eval net, and this set of parameters is fixed for a period of time before being replaced by a new parameter of eval net. Eval net is increasing, so it is a trainable network, but target net is not trainable.[15]

3.3. Optimization of DQN by Dense Network and Squeeze Excitation Network

Since Res-Net was put forward, the variety of Res-Net network emerged in an endless stream, each has its own characteristics, the network performance has also been improved. Dense-Net[3], proposed in 2017, absorbs the best parts of Res-Net and does more innovative work on it to further improve network performance.

In general, CNN network shall reduce the size of feature map through Pooling or stride>1 Conv, while the dense connection mode of Dense-Net shall maintain the same size of feature map. In order to solve this problem, the dense block with transition structure is used in the Dense-Net, in which dense block is a module containing many layers, each layer has the same size feature map, and the layers are connected in a dense way. Transition module connects two adjacent dense block and reduces the size of feature map by pooling.

Dense-Net is a neural network with denser connections than Res-Net. In this network, the outputs of previous layers are made best use by the other layers after the previous layers. The structure inside the block is as figure 3. The Dense block structure is similar to Res-Net block: Batch Normalization (BN), ReLU and Conv. A Dense-Net is made up of few such blocks. The transition layers are between each dense block which consist of BN, Conv and Average-Pooling.

![Figure 3. Structure of Dense Block.](image)

Although Dense-Net uses dense connections, in fact it is more efficient than other networks. The key is to reduce the amount of computing at each layer of the network and to reuse the features. Dense-Net has a very good anti-overfitting performance, especially suitable for the application of relatively scarce training data. Dense-Net improves parameter utilization through more connections and directional changes reduce the total number of parameters required. For the specific indoor environment with fewer environmental characteristics, the network can use fewer parameters, and the training effect using Dense-Net is better than that using Res-Net. This is evident in Dense-Net's presentation in the paper of a CIFAR dataset that does not do data augmentation. For example, without data enhancement for CIFAR100, the previous best result was 28.20% error rate, while Dense-Net can improve this result to 19.64%[16]

Besides the Dense-Net, there are many ways to improve the performance of the network from the spatial dimension. For example, the inception structure embedding multi-scale information and aggregating a variety of features on different receptive fields to obtain performance gains; the context information in the space is considered in the inside-outside network. And then there's the introduction of the attention mechanism into the spatial dimension. Squeeze-and-Excitation Networks (SE-Net),
proposed by Hu Jie's team in 2019 [17], instead of introducing a new spatial dimension for the fusion of feature channels, a new "feature recalibration" strategy is adopted. Specifically, it is to automatically obtain the importance of each feature channel through learning, and then promote the useful features according to this importance and suppress the features that are not useful for the current task. In a nutshell, the core of the SE-Net is channel attention which is shown below.

![Figure 4. The Structure of Se-Net.](image)

The SE block is not the complete network structure, it is simple to construct, easy to deploy, does not require the introduction of new functions or layers, and can be embedded in almost any network architecture today. In addition, it has good characteristics in terms of model and computational complexity. According to the data in the paper mentioned above, the accuracy of SE-Res-Nets at all depths far exceeds that of their non-SE counterparts, indicating that SE modules can bring performance gains to the network regardless of the depth of the network.

3.4. Proposed Method

This paper used the information of camera and LIDAR to accomplish the task of navigation. In order to process the images, Bilinear Interpolation was adapted to change the size of images without pooling or convolution. The information of LIDAR was processed by LSTM and then merged with the processed information of image.

The whole process of the algorithm of this paper shows below:

![Figure 5. The Process of the SE-Dense-DQN Algorithm.](image)

With regard to Dense-Net, assume that the number of channels in the feature graph obtained by connecting Dense block on the Transition layer is \( k \). The Transition layer can generate \( m_k \) features...
through the convolution layer, where m is the compression rate (0<\(m\)<1). When \(m=1\), the number of features does not change through the Transition layer, that is, there is no compression. When the compression coefficient is less than 1, this structure is called Dense-Net-BC. The Dense Block structure that uses a bottleneck layer and the Transition combined structure with a compression coefficient less than 1 are called Dense-Net-BC. We can get the structure of Dense-Net-BC by taking SE module as a bottleneck. The introduction method is similar to the SE-Res-Net which has been tested by the author who proposed the SE structure [17]. The framework of the Q Network in this paper is shown in figure 6.

4. Experiments and Results

This paper built the network by the PyTorch and carried out the experiment with the open-source project PX4 Autopilot and Robot Operating System. To visualize the results of simulation, Gazebo was also used in the experiment. Since the dataset is not needed for DQN, the data that to be trained is accumulated during the random exploration of agent, this paper would not provide the dataset. The 3D simulation environment is an Open-Source project called XTDrone which is shown below. This paper carried out the experiment with the model of UAV with LIDAR and camera in the complex scene, and set the start point and the target point in the scene. When the UAV starts its work, it explored the scene randomly and then used the data accumulated before to train the Q network.

Since the DQN is similar to Reinforcing Learning that tends to choose maximum reward action, the reward function was set as follows [6]:

\[
R(s_t, a_t) = \begin{cases} 
\gamma \cdot l_i & l_i \geq 0.1m \\
-20 & l_i < 0.1m
\end{cases}
\]
$s_t$ is the present state, $a_t$ is the possible action may be taken, the $l_i (0 < l_i < 8)$ stands for the present distance of the barriers and the agent, $l_{tt}$ represents the distance between the agent and the target and $\gamma$ is the parameter of the reward gain.

To judge the performance of the network of this paper, the first metric is the arrival score [6][3]. This paper changed the targets to do tests among different networks like DQN and DDQN (Double DQN) and calculate the arrival score. The formula of mean arrival score is calculated by reward value:

$$\delta_n = \frac{1}{N_{total}} \cdot \sum_{i=1}^{N_{total}} R_i$$

Another metric that could qualify the network is the convergence rate, which reflects the efficiency of the network with the value of loss function during the increasing of steps.

The final result of simulation and comparison with other networks are shown below:

5. Conclusion

In order to improve the indoor navigation performance of UAVs, this paper proposed the method of fuse information of LIDAR and images to make sure better accuracy. The shift-invariance and the more information merged makes the DQN more adaptable. Results of arrival score and the convergence rate shows that the SE-Dense-DQN is better than other classic DQN. The better performance assures the better safety of UAVs adopted this algorithm.
However, this method adopted by paper consumes amounts of computing resources which make the onboard computer difficult to handle the data. The optimization of the method needs to be conduct in the future.

References

[1] Lázaro M T, Capobianco R and Grisetti G 2018 Efficient Long-term Mapping in Dynamic Environments 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Madrid, Spain pp 153-160

[2] Chen X, Zhang H, Lu H, Xiao J, Qiu Q and Li Y 2017 Robust SLAM system based on monocular vision and LiDAR for robotic urban search and rescue 2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR), Shanghai, China pp 41-47

[3] Sun J, Qian Z, Wang X and Wang X 2020 ES-DQN-Based Vertical Handoff Algorithm for Heterogeneous Wireless Networks IEEE Wireless Communications Letters vol 9 pp 1327-1330

[4] Yue P, Xin J, Zhao H, Liu D, Shan M and Zhang J 2019 Experimental Research on Deep Reinforcement Learning in Autonomous navigation of Mobile Robot 2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA), Xi'an, China pp 1612-1616

[5] Gui B, Qian K, Chen S, Jing X and Ma X 2020 Knowledge Induced Deep Q-Network for Robot Push and Grasp Manipulation Skills Learning 2020 Chinese Automation Congress (CAC), Shanghai, China pp 4078-4083

[6] Nilwong S and Capi G 2020 Outdoor Robot Navigation System using Game-Based DQN and Augmented Reality 2020 17th International Conference on Ubiquitous Robots (UR), Kyoto, Japan pp 74-80

[7] Li Z and Guo G 2020 Research on Optimization Strategy Based on Improved DDPG Algorithm in WSN Proceedings of the 2020 4th International Conference on Digital Signal Processing (ICDSP 2020) pp 358–362

[8] Sarthak B and Sujit P B UAV 2020 Target Tracking in Urban Environments Using Deep Reinforcement Learning arXiv:2007.10934v1

[9] Duc N T, Hai Q T, Van D N, Trong T H and Trong T D 2019 An approach for UAV indoor obstacle avoidance based on AI technique with ensemble of ResNet8 and Res-DQN 2019 6th NAFOSTED Conference on Information and Computer Science (NICS)

[10] Huang H, Yang Y, Wang H, Ding Z, Sari H and Adachi F 2020 Deep Reinforcement Learning for UAV Navigation Through Massive MIMO Technique IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY vol 69

[11] Zhang R 2019 Making Convolutional Networks Shift-Invariant Again Proceedings of the 36th International Conference on Machine Learning Long Beach, California p 97

[12] Singla A Padakandla S and Bhatnagar S 2021 Memory-Based Deep Reinforcement Learning for Obstacle Avoidance in UAV With Limited Environment Knowledge IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS vol 22 NO. 1

[13] Sutton R S and Andrew G B 1998 Reinforcement Learning: An Introduction Adaptive Computation and Machine Learning series, MIT Press (Bradford Book) Cambridge, Mass vol 17 pp 229-235

[14] Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D and Riedmiller M 2013 Playing Atari with Deep Reinforcement Learning. arXiv:1312.5602

[15] Yang Y, Wang Y, Liu K, Zhang N, Gu Z and Zhang Q Deep Reinforcement Learning Based Online Network Selection in CRNs With Multiple Primary Networks IEEE Transactions on Industrial Informatics 2020

[16] Huang G, Liu Z, Maaten L V and Weinberger K Q 2018 Densely Connected Convolutional Networks arXiv:1608.06993v5

[17] Zhao C, Shuai R, Ma L, Liu W, Hu D and Wu M Dermoscopy Image Classification Based on StyleGAN and DenseNet201 IEEE Access 2021

[18] Huang J, Yang Y, Wang H, Ding Z, Sari H and Adachi F Deep Reinforcement Learning for UAV Navigation Through Massive MIMO Technique IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY 2020